现代数学丛书

袁亚湘 著

非线性规划数值方法

NUMERICAL METHODS FOR NONLINEAR PROGRAMMING

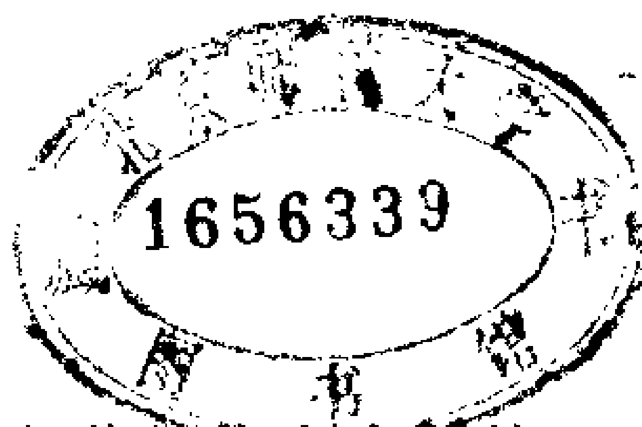
YA-XIANG YUAN

上海科学技术出版社

·现代数学丛书·

非线性规划数值方法

袁亚湘　著



上海科学技术出版社

责任编辑　赵序明

现代数学丛书
非线性规划数值方法
袁亚湘　著
上海科学技术出版社出版、发行
（上海瑞金二路450号）
新华书店上海发行所经销　商务印书馆上海印刷厂印刷
开本 787×1092　1/20　印张 18　插页 4　字数 225 000
1993年8月第1版　1993年8月第1次印刷
印数　1—1,500
ISBN 7-5323-3012-5/O·165
定价：10.20元

（沪）新登字108号

内 容 提 要

本书介绍非线性规划的主要数值方法与理论。书中许多章节是基于作者的博士论文及作者近年来的科研成果。它包括无约束优化、二次规划、非线性约束优化以及非光滑优化等方法的最新成果。书中着重介绍了近年来的新方法及结果，如：信赖域方法、既约海色阵方法的收敛特征、拟牛顿法的收敛性、Barzilai–Borwein 梯度法的收敛速度等等。书中也介绍了非线性规划的主要经典方法，如：梯度法、共轭梯度法、牛顿法、拟牛顿法、逐步二次规划法、既约梯度法等等以及这些常用方法的最新进展。

本书是一部学术性专著，介绍了当今国际上非线性规划前沿的研究成果，引用了大量近期的、较新的参考文献，以方便希望进一步研究的同行们。

本书可作为广大非线性规划研究人员以及从事实际应用的工程技术人员的参考书，也可作为研究生和大学高年级学生的教材。

Modern Mathematics Series

NUMERICAL METHODS FOR NONLINEAR PROGRAMMING

Ya-xiang Yuan

Shanghai Scientific & Technical Publishers

Numerical Methods For Nonlinear Programming

Abstract

This book discusses numerical methods for nonlinear programming. Its contents include recent results on unconstrained optimization, quadratic programming, nonlinearly constrained optimization and nonsmooth optimization. Newly developed numerical methods such as trust region algorithms, interior point methods and sequential quadratic programming methods are discussed. Many recent results such as one step fast one step slow convergence of reduced Hessian algorithms, super-linear convergence of Barzilai-Borwein gradient method, and global convergence of quasi-Newton algorithms are given. Classical methods including steepest descent method, Newton's method, quasi-Newton method, reduced gradient method are discussed and recent advances of these methods are also presented.

出版说明

从六十年代起，由华罗庚教授任主编的《现代数学丛书》编辑委员会曾组织编著，并由我社出版了多部具有很高水平的数学学术专著，有几部专著并已在国外出了外文版，受到国内外数学界和广大读者的高度重视，获得了很高的评价．原编委会中华罗庚、关肇直、吴新谋三位教授虽已先后逝世，但他们为本《丛书》所作出的贡献迄今仍为人们所敬仰、怀念．由于某些客观原因，《现代数学丛书》的出版工作曾一度停顿．

为了适应现代数学的迅速发展，更好地反映我国数学家近几年的优秀研究成果，必须大力加强《现代数学丛书》的规划、编辑、出版工作．充实编委会的力量．考虑到不少编委年事已高，经向原编委会中大部分同志及数学界有关专家广泛征求意见后，于1990年对编委会作了调整，补充了一些著名的中年数学家和学科带头人，建立了新的编委会，并进一步明确了本丛书的宗旨．

《现代数学丛书》新的编辑委员会由苏步青教授任名誉主编、谷超豪教授任主编，十八位著名数学家任委员．编委会负责推荐（或审定）选题和作者，主持书稿的审核等工作．

《现代数学丛书》的宗旨是：向国内外介绍我国比较成熟的、对学科发展方向有引导作用的、国内第一流水平的数学研究成果，反映我国数学研究的特色和优势，扩大我国数学研究成果的影响，促进学科的发展和国内外的学术交流．

为了实现上述宗旨，本丛书将陆续组织出版在基础数学、应用数学和计算数学方面处于学科发展前沿、有创见且具有系统完整

研究成果的现代数学学术专著.

为出版好《现代数学丛书》，我们热切地期望着数学界各位专家的大力支持和悉心指导，并欢迎广大读者提出宝贵的建议和意见。

上海科学技术出版社

前　言

非线性规划是计算数学和运筹学交叉的学科．它在工程、国防、经济等许多重要领域有着广泛的应用．近二十多年来，非线性规划一直是一个十分活跃的研究领域．本书总结了作者近年来的科研结果（研究工作曾得到国家自然科学基金资助），且系统地介绍了国际上主要的非线性规划计算方法及其理论．

非线性规划从七十年代起就是数学规划中最受重视的分支之一．特别是变尺度算法的兴起使国际上涌现出来如 Powell、Fletcher、Dennis 等一批著名的非线性规划专家．近年来，求解大规模优化问题的 SQP 方法的成功应用，使得非线性规划计算方法成为大规模科研与工程计算的一个重要方向．八十年代初期兴起的信赖域法已是非线性规划的一个十分热门的研究课题，它进一步促进了非线性规划的研究与应用．

我国关于非线性规划的研究起始于 1970 年．当时，由华罗庚教授领导的中国科学院数学研究所运筹研究室对优化方法进行研究、应用和推广．尔后，吴方、桂湘云、席少霖、郑权、俞文魮等教授对非线性规划进行了大量的科研．目前我们已有一大批从事非线性规划研究的专业人员．在应用方面，继华罗庚教授推广优选法之后，钱令希、张可村、周济等教授又做了许多工作，已取得不少优秀的成果．

本书的许多章节是基于作者的博士论文以及作者近年来的科研结果．第二章给出的一个不需要计算二阶导数的平方收敛方法；第三章给出的 Barzilai-Borwein 梯度法的超线性收敛性和关于共轭梯度法的线性收敛速度的上界估计；第四章给出的精确搜索下变尺度方法的最小收敛阶、Broyden 族自对偶修正公式的导

出以及整个 Broyden 凸族(DFP 方法除外)的整体收敛性；第八章的求解线性约束问题的信赖域法；第九章提出的一个 SQP 方法(第 9.3 节)和一个信赖域方法(第 9.5 节)以及关于既约海色阵方法的一快一慢收敛特征(第 9.5 节)以及第十章给出的关于非光滑优化信赖域法的一系列结果都是作者或作者与 Powell、Byrd 和 Nocedal 等教授合作完成的研究结果.

本书是一本学术性专著，介绍了当今国际上非线性规划前沿的研究成果.引用了大量的较新的参考文献，以提供希望进一步研究的同行们参考. 本书可作为广大非线性规划研究人员以及从事实际应用的工程技术人员的参考书，也可作为研究生或大学高年级学生的教材.

由于作者长期在国外，本书对国内的优秀工作介绍不多，但愿在以后再版时能弥补这一不足. 另外，限于作者的水平，书中可能有不妥或错误之处，敬请广大读者批评指正.

我想借此机会对我的博士论文指导导师、英国皇家学会会员、剑桥大学的 M. J. D. Powell 教授表示衷心的感谢，我在非线性规划方面所取得的科研成果都和他的指导与鼓励是分不开的. 我非常感谢中国科学院学部委员、中国科学院计算中心名誉主任冯康教授多年来对我的关心和培养；感谢中国科学院计算中心石钟慈、赵风治等老师对我写这本书的鼓励. 本书的部分书稿是作者于 1990 年 9 月至 1991 年 8 月访问美国完成的，所以我还感谢美国科罗拉多大学 Boulder 分校的 Richavd H. Byrd 教授和美国西北大学的 Jorge Nocedal 教授在我访问他们学校期间提供的良好工作条件以及友好的很有成效的科研合作. 北京计算机学院席少霖教授审阅了本书，提出了宝贵的意见和建议. 我向他表示深切的谢意. 最后，我感谢我的爱人也是我的同事王慧娟多年来给我的鼓励、帮助和支持.

袁亚湘

1992 年 3 月于

中国科学院计算中心

目　录

CONTENTS

第 1 章

绪　论

§1.1 问　题

非线性规划(也称非线性优化)问题是求一个定义在 n 维空间的单值函数的极值问题，函数的自变量可能受限制于有限个不等式或等式约束. 通常有如下形式:

$$\min_{x\in \mathbf{R}^n} f(x); \tag{1.1.1}$$

$$\text{s.t. } C_i(x)=0 \quad (i=1,\ 2,\ \cdots,\ m_e); \tag{1.1.2}$$

$$C_i(x)\geqslant 0 \quad (i=m_e+1,\ \cdots,\ m). \tag{1.1.3}$$

$f(x)$ 及 $C_i(x)$ $(i=1,\ 2,\ \cdots,\ m)$ 都是定义在 $\mathbf{R}^n$ 上的函数，其中至少有一个为非线性函数. m 和 m_e 是两个非负整数且满足 $m\geqslant m_e$. s. t. 是英文subject to(满足于)的缩写. $f(x)$称为目标函数，$C_i(x)$ $(i=1,\ 2,\ \cdots,\ m)$ 称为约束函数，(1.1.2)和(1.1.3)称为约束条件. 满足约束条件的点称为可行点，所有可行点的集合称为可行域，记为 X. 显然，

$$X=\{x\,|\,C_i(x)=0,\ i=1,\ \cdots,\ m_e;\ C_i(x)\geqslant 0,\ i=m_e+1,\ \cdots,\ m\}. \tag{1.1.4}$$

利用定义(1.1.4)，我们可将问题(1.1.1)～(1.1.3)写成如下形式:

$$\min_{x\in X} f(x). \tag{1.1.5}$$

当 $m=0$ 时，问题(1.1.1)～(1.1.3)是无约束优化问题，否则，是约束优化问题. 在约束优化问题中，如果只有等式约束，即 $m=m_e>0$，则称为等式约束优化问题. 另一种特殊情形是所有的约

束函数都是线性函数，这时我们称问题(1.1.1)～(1.1.3)是一个线性约束优化问题.

非线性规划这一名词是由 Kuhn 和 Tucker(1950) 最先提出来的. 但非线性规划问题早已有之，例如 Sylvester(1857)提出的在平面上找一个包含 N 个给定点的最小圆的问题，它可写成如下极小化问题：

$$\min_{x\in\mathbf{R}^2} f(x) \triangleq \max_{1\leqslant i\leqslant N} \|x-x_i\|_2, \tag{1.1.6}$$

其中 $\|\cdot\|_2$ 是欧氏范数，$x_i(i=1, 2, \cdots, N)$是平面上的 N 个给定点. 显然，(1.1.6)的解 x^* 即是所求圆的中心，$f(x^*)$是所求圆的半径.

非线性曲线拟合问题通常可写成如下最小二乘的形式，即

$$\min_{x\in\mathbf{R}^n} f(x) \triangleq \sum_{i=1}^{N}[\varphi(x, t_i)-y_i]^2, \tag{1.1.7}$$

其中 $t_i\in\mathbf{R}^m$, $y_i\in\mathbf{R}(i=1, 2, \cdots, N)$ 是实验值, x 是特定参数, $\varphi(x, t)$是一非线性函数，它是取定的拟合曲线的基函数. 设 x^* 是(1.1.7)的解，则

$$y(t)=\varphi(x^*, t) \tag{1.1.8}$$

就是通过最小二乘对数据$(t_i, y_i)(i=1, 2, \cdots, N)$的拟合曲线.

非线性优化问题广泛见之于工程、国防、经济、管理等许多重要领域，在结构设计、化学反应设计、电力分配、石油开采等方面都有直接的应用. 非线性规划计算方法还和计算数学中的数值逼近、常微分方程数值解、偏微分方程中的变分原理、微分方程反演以及非线性代数方程数值解等分支和问题有交叉和应用.

定义 1.1.1 设 $x^*\in X$，如果对某一 $\delta>0$ 有

$$f(x)\geqslant f(x^*),\ x\in X\cap B(x^*, \delta)\setminus\{x^*\}, \tag{1.1.9}$$

则称 x^* 是问题(1.1.1)～(1.1.3)的局部极小值，其中

$$B(x^*, \delta)=\{x \mid \|x-x^*\|_2\leqslant\delta\}. \tag{1.1.10}$$

如果(1.1.9)式对严格不等号“$>$”成立，则称 x^* 是局部严格极小值.

定义 1.1.2 设 $x^*\in X$，如果

$$f(x)\geqslant f(x^*),\ x\in X\setminus\{x^*\},\tag{1.1.11}$$

则称 x^* 是问题(1.1.1)～(1.1.3)的全局极小值. 如果(1.1.11)式对严格不等号“>”成立,则称 x^* 是全局严格极小值.

全局极小值也被称作总体极小值. 从定义可知, 全局(严格)极小值必定是局部(严格)极小值.

定义 1.1.3 对任何 $x\in \mathbf{R}^n$, 我们称集合

$$A(x)=E\cup I(x)\tag{1.1.12}$$

是在 x 点的积极集合, 其中 $E=\{1,\ 2,\ \cdots,\ m_e\}$, $I=\{m_e+1,\ \cdots,\ m\}$ 以及

$$I(x)=\{i\,|\,c_i(x)\leqslant 0,\ i\in I\}.\tag{1.1.13}$$

我们称 $c_i(x)$ $(i\in A(x))$是在 x 点的积极约束, 称$c_i(x)$ $(i\notin A(x))$是在 x 点的非积极约束.

§1.2 最优性条件

对于给定的 $x^*\in X$, 在一般情况下, 集合 $X\cup B(x^*,\ \delta)\setminus\{x^*\}$ 中有无穷多个点. 所以, 通过验证(1.1.9)来判别 x^* 是否是优化问题(1.1.1)～(1.1.3)的解几乎是不可能的. 最优性条件就是在一定条件下和(1.1.9)等价的条件, 这些条件仅依赖于目标函数和约束函数在 x^* 点的性质.

对于无约束问题, 显然有以下定理:

定理 1.2.1 假设 x^* 是无约束问题(1.1.1)的局部极小点, $f(x)$在 x^* 处可微, 则必有

$$\nabla f(x^*)=0;\tag{1.2.1}$$

如果 $f(x)$在 x^* 处二次可微, 则必有

$$d^T\nabla^2 f(x^*)d\geqslant 0\tag{1.2.2}$$

对任何 $d\in \mathbf{R}^n$ 都成立.

定理 1.2.1 是一个必要性定理, 它给出在解处函数必需满足的条件. 满足(1.2.1)的点称为函数 $f(x)$的稳定点. 从上述可知, 只要目标函数可微, 局部极小点必是稳定点. 下面的定理给出稳

定点是局部极小点的充分条件.

定理 1.2.2 假设 $f(x)$ 在 x^* 处二次连续可微, 如果 x^* 是 $f(x)$ 的稳定点且对任何非零向量 $d\in\mathbf{R}^n$ 有

$$d^T\nabla^2 f(x^*)d>0, \tag{1.2.3}$$

则 x^* 是无约束问题(1.1.1)的局部严格极小点.

以上两个定理的证明十分简单, 只需利用 Taylor 展开即可, 故从略.

现在我们考虑约束问题. 显然, 可行域上一个点是否为局部极小点取决于目标函数在这一点以及在该点附近的可行点的值. 所以我们有必要给出可行方向的定义.

定义 1.2.3 设 $x^*\in X$, $d\in\mathbf{R}^n$, 如果存在 $\delta>0$ 使得

$$x^*+td\in X,\ \forall t\in[0,\ \delta], \tag{1.2.4}$$

则称 d 是 X 在 x^* 处的可行方向. X 在 x^* 处的所有可行方向的集合记为 $FD(x^*,\ X)$.

定义 1.2.4 设 $x^*\in X$, 我们称集合

$$\begin{aligned}\mathrm{LFD}(x^*,\ X)=\{d\ |\ &d^T\nabla c_i(x^*)=0,\ i\in E,\\ &d^T\nabla c_i(x^*)\geqslant 0,\ i\in I(x^*)\}\end{aligned} \tag{1.2.5}$$

为 X 在 x^* 处的线性化可行方向集合; 如果 $d\in\mathrm{LFD}(x^*,\ X)$, 则称 d 是 X 在 x^* 处的线性化可行方向.

定义 1.2.5 设 $x^*\in X$, 我们称集合

$$\mathrm{SFD}(x^*,\ X)=\{d\ |\ d_k\to d,\ x^*+\delta_k d_k\in X,\ \delta_k>0,\ \delta_k\to 0\} \tag{1.2.6}$$

为 X 在 x^* 处的序列可行方向集合; 如果 $d\in\mathrm{SFD}(x^*,\ X)$, 则称 d 是 X 在 x^* 处的序列可行方向.

根据定义, 下列引理显然成立:

引理 1.2.6 如果所有的约束函数都在 $x^*\in X$ 处可微, 则有

$$\mathrm{FD}(x^*,\ X)\subseteq\mathrm{SFD}(x^*,\ X)\subseteq\mathrm{LFD}(x^*,\ X). \tag{1.2.7}$$

引理 1.2.7 设 x^* 是问题(1.1.1)~(1.1.3)的局部极小点, 则必有

$$d^T\nabla f(x^*)\geqslant 0,\ \forall d\in\mathrm{SFD}(x^*,\ X). \tag{1.2.8}$$

证明　由定义 1.2.5, 对任何 $d\in SFD(x^*, X)$, 存在 $\delta_k\to 0$, $d_k\to d$. 使得 $x^*+\delta_k d_k\in X$. 因为 x^* 是局部极小点, 所以对充分大的 k, 必有

$$f(x^*+\delta_k d_k)\geqslant f(x^*). \tag{1.2.9}$$

由于 $f(x)$ 在 x^* 处可微, 从(1.2.9)可知

$$\delta_k d^T\nabla f(x^*)+o(\delta_k)\geqslant 0. \tag{1.2.10}$$

因为 $\delta_k>0$ 且 $\delta_k\to 0$, 在上式两端除以 δ_k 后, 令 $k\to\infty$, 即可得(1.2.8). ▌

下面的引理是由 Farkas(1902)给出的, 故称为 Farkas 引理. 由于它在形式上是线性方程组及线性不等式组(1.2.11)～(1.2.13)和线性表达式(1.2.14)这两者中必有一个且只有一个成立, 所以该引理也称为择一性引理.

引理 1.2.8　设 l, l' 是两个非负整数, a_0, $a_i(i=1, 2, \cdots, l)$ 和 $b_i(i=1, 2, \cdots, l')$ 是 $\mathbf{R}^n$ 中的向量, 则线性方程组及不等式组

$$d^T a_i=0\quad (i=1, 2, \cdots, l), \tag{1.2.11}$$

$$d^T b_i\geqslant 0\quad (i=1, 2, \cdots, l'), \tag{1.2.12}$$

$$d^T a_0<0 \tag{1.2.13}$$

无解当且仅当存在实数 $\lambda_i(i=1, 2, \cdots, l)$ 和非负实数 $\mu_i(i=1, 2, \cdots, l')$, 使得

$$a_0=\sum_{i=1}^{l}\lambda_i a_i+\sum_{i=1}^{l'}\mu_i b_i. \tag{1.2.14}$$

证明　假定(1.2.14)成立, 且 $\mu_i\geqslant 0(i=1, \cdots, l')$, 如果(1.2.11)和(1.2.12)成立, 则

$$d^T a_0=\sum_{i=1}^{l}\lambda_i d^T a_i+\sum_{i=1}^{l'}\mu_i d^T b_i\geqslant 0. \tag{1.2.15}$$

故知(1.2.13)不成立. 从而(1.2.11)～(1.2.13)无解.

假定(1.2.14)不可能成立. 定义集合

$$S=\left\{a;\ \middle|\ a=\sum_{i=1}^{l}\lambda_i a_i+\sum_{i=1}^{l'}\mu_i b_i,\ \lambda_i\in\mathbf{R}, \mu_i\geqslant 0\right\}. \tag{1.2.16}$$

显然 S 是 $\mathbf{R}^n$ 中的一个闭凸锥. 由于 $a_0\notin S$, 根据泛函分析中的凸集分离定理, 必存在 $d\in\mathbf{R}^n$ 和 $\alpha\in\mathbf{R}$, 使得

$$d^T a_0 < \alpha < d^T a, \ \forall a \in S. \tag{1.3.17}$$

由于 $0 \in S$，所以

$$d^T a_0 < 0. \tag{1.2.18}$$

对任何 $\lambda > 0$，均有 $\lambda b_i \in S$. 从而

$$\lambda d^T b_i > \alpha. \tag{1.2.19}$$

在上面不等式两边同除以 λ，然后令 $\lambda \to +\infty$，即得

$$d^T b_i \geqslant 0 \quad (i=1, 2, \cdots, l'). \tag{1.2.20}$$

对于任何 $\lambda > 0$，均有 $\lambda a_i \in S$，$-\lambda a_i \in S$，同上可以证 $d^T a_i \geqslant 0$ 和 $d^T(-a_i) \geqslant 0$. 故知

$$d^T a_i = 0 \quad (i=1, 2, \cdots, l). \tag{1.2.21}$$

所以(1.2.11)～(1.2.13)有解. ▮

利用 Karkas 引理和引理 1.2.7，可证明一阶优化条件：

定理 1.2.9(**Kuhn-Tucker 定理**) 设 x^* 是 (1.1.1)～(1.1.3)的一个局部极小点，如果

$$\mathrm{SFD}(x^*, X) = \mathrm{LFD}(x^*, X), \tag{1.2.22}$$

则存在 $\lambda_i^*(i=1, 2, \cdots, m)$，使得

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i^* \nabla C_i(x^*), \tag{1.2.23}$$

$$\lambda_i^* \geqslant 0, \ \lambda_i^* C_i(x^*) = 0 (i \in I). \tag{1.2.24}$$

证明 由引理 1.2.7 和式(1.2.22)即知线性方程及不等式组

$$d^T \nabla C_i(x^*) = 0 \quad (i \in E), \tag{1.2.25}$$

$$d^T \nabla C_i(x^*) \geqslant 0 \quad (i \in I(x^*)), \tag{1.2.26}$$

$$d^T \nabla f(x^*) < 0 \tag{1.2.27}$$

无解. 从引理 1.2.8 可知，存在 $\lambda_i^*(i \in E)$ 及 $\lambda_i^* \geqslant 0 (i \in I(x^*))$ 使得

$$\nabla f(x^*) = \sum_{i \in E} \lambda_i^* \nabla C_i(x^*) + \sum_{i \in I(x^*)} \lambda_i^* \nabla C_i(x^*). \tag{1.2.28}$$

令 $\lambda_i^* = 0 (i \in I \setminus I(x^*))$，即知(1.2.23)～(1.2.24)成立. ▮

与(1.2.23)式有密切联系的一个函数是

$$L(x, \lambda) = f(x) - \lambda^T C(x) = f(x) - \sum_{i=1}^{m} \lambda_i C_i(x). \tag{1.2.29}$$

其中 $\lambda=(\lambda_1, \cdots, \lambda_m)^T\in\mathbf{R}^m$, $C(x)=(C_1(x), \cdots, C_m(x))^T$. 由于这一函数的思想可追溯到 Lagrange(1760～1761), 故它被称为 Lagrange 函数; $\lambda_i(i=1, \cdots, m)$被称为 Lagrange 乘子. 由于定理 1.2.9 是 Kuhn 和 Tucker (1951)给出的, 我们给出如下定义:

定义 1.2.10 如果 $x^*\in X$ 且存在 $\lambda^*=(\lambda_1^*, \cdots, \lambda_m^*)\in\mathbf{R}^m$ 满足 (1.2.23)～(1.2.24), 则称 x^* 是问题 (1.1.1)～(1.1.3) 的 Kuhn-Tucker 点(简称 K-T 点).

因为 Karush(1939)也类似地考虑了约束优化的最优性条件, 所以也有人将定理 1.2.9 称为 Karush-Kuhn-Tucker 定理, 把 K-T 点称为 K-K-T 点.

定理 1.2.9 中的条件(1.2.22)称为约束规范条件. Mangasarian 和 Fromowitz(1967)提出另一个约束规范条件:

$$\nabla C_1(x^*), \cdots, \nabla C_{m_e}(x^*)\text{线性无关}, \tag{1.2.30}$$

$$S^*=\{d\,|\,d^T\nabla C_i(x^*)=0,\ i\in E,\ d^T\nabla C_i(x^*)>0,\ i\in I(x^*)\}\text{非空}, \tag{1.2.31}$$

并且给出了下面的定理:

定理 1.2.11 在定理 1.2.9 中, 将条件 (1.2.22) 换成 (1.2.30)～(1.2.31), 必要性条件(1.2.23)～(1.2.24)仍然成立.

证明 $\forall d\in S^*$, $d\neq0$, 必存在 $d_i(i=1, \cdots, m-m_e-1)$组成 span$\{\nabla C_1(x^*), \cdots, \nabla C_{m_e}(x^*), d\}$的法空间的一规正交基. 考虑带参数的非线性代数方程组:

$$C_i(x)=0 \quad (i=1, \cdots, m_e), \tag{1.2.32}$$

$$d_i^T(x-x^*)=0 \quad (i=1, \cdots, m-m_e-1), \tag{1.2.33}$$

$$d^T(x-x^*)-\theta=0. \tag{1.2.34}$$

由于该方程组在 $x=x^*$ 处的 Jacobi 矩阵是非奇异的, 根据隐函数定理, 对充分小的 θ, 必存在解 $x=x(\theta)$, 且满足

$$x'(\theta)|_{\theta=0}=d. \tag{1.2.35}$$

$\forall i\in I(x^*)$, 有 $d^T\nabla C_i(x^*)>0$, 故当 $\theta>0$ 充分小时, 必有

$$C_i(x(\theta))>0.$$

所以 $x(\theta)\in X(\theta>0,\ \theta$ 充分小). 取 $\theta_k>0$, $\theta_k\to0(k\to\infty)$, 且

$x(\theta_k)\in X$，由定义 1.2.5 和关系式(1.2.35)即知

$$d\in \mathrm{SFD}(x^*,\ X). \tag{1.2.36}$$

由于 $d\in S^*$ 的任意性，我们可知

$$S^*\subseteq \mathrm{SFD}(x^*,\ X). \tag{1.2.37}$$

从引理 1.2.7 以及上式可得

$$d^T f(x^*)\geqslant 0,\ \forall d\in S^*. \tag{1.2.38}$$

故必有

$$d^T f(x^*)\geqslant 0,\ \forall d\in \mathrm{Cl}S^*, \tag{1.2.39}$$

其中 $\mathrm{Cl}S^*$ 表示 S^* 的闭包. 由于 S^* 非空，根据 Rockafellar(1970) 中的定理 6.5 可知

$$\mathrm{Cl}S^*=\{d\,|\,d^T\nabla C_i(x^*)=0,\ i\in E,\ d^T\nabla C_i(x^*)\geqslant 0,\ i\in I(x^*)\}. \tag{1.2.40}$$

式 (1.2.39)～(1.2.40) 表明线性方程及不等式组 (1.2.25)～(1.2.27) 无解. 故知存在 $\lambda_i^*(i=1,\ 2,\ \cdots,\ m)$ 使得必要性条件 (1.2.23)～(1.2.24) 成立. ▮

一个比(1.2.30)～(1.2.31)更强的约束规范条件是

$$\nabla C_i(x^*)\quad (i\in E\cup I(x^*))\text{线性无关}. \tag{1.2.41}$$

显然，当 $I(x^*)=\varnothing$ 时，(1.2.41)和(1.2.30)～(1.2.31)等价；当 $I(x^*)\neq\varnothing$ 时，利用 Farkas 引理不难证明当 $\nabla C_i(x^*)(i\in E\cup I(x^*))$ 线性无关时，S^* 一定非空. 从而有以下定理.

定理 1.2.12 在定理 1.2.9 中，将条件 (1.2.22) 换成 (1.2.41)，必要性条件(1.2.23)～(1.2.24)仍然成立.

因为约束规范条件(1.2.41)易于检验，所以定理 1.2.12 是一个最常见的也是最有用的关于一阶最优性条件的结果.

下面两个定理是众所周知的(见 Fletcher, 1981)，此二定理给出了局部极小点的二阶必要条件和充分条件.

定理 1.2.13 假定 $f(x)$、$C_i(x)(i=1,\ \cdots,\ m)$都在 x^* 附近两次连续可微，如果 x^* 是问题(1.1.1)～(1.1.3)的局部极小点且 $\nabla C_i(x^*)(i\in E\cup I(x^*))$ 线性无关，则必存在 λ^*，使得 (1.2.23)～(1.2.24)成立，且

$$d^T\nabla_{xx}L(x^*,\ \lambda^*)d\geqslant 0,\ \forall d\in S(x^*). \qquad (1.2.42)$$

其中

$$S(x^*)=\{d\mid d\neq 0,\ d^T\nabla C_i(x^*)=0.\ 如果\ i\in E,$$
$$\lambda_i^*>0,\ d^T\nabla C_i(x^*)\geqslant 0,\ \forall i\in I(x^*)\}. \qquad (1.2.43)$$

定理1.2.14 假定$f(x)$, $C_i(x)(i=1,\ \cdots,\ m)$都在$x^*\in X$附近两次连续可微，如果存在λ^*使得(1.2.23)～(1.2.24)成立，且

$$d^T\nabla_{xx}L(x^*,\ \lambda^*)d>0,\ \forall d\in S(x^*), \qquad (1.2.44)$$

则x^*必定是问题(1.1.1)～(1.1.3)的局部严格极小点.

在更弱的约束规范条件下，甚至没有任何约束规范条件也可以给出最优性条件(如 Guignard, 1982; Mc Cormick, 1983). 另外，高阶的最优性条件也可给出(如 Lempio 和 Zowe, 1982).

如果$f(x)$、$C_i(x)(i=1,2,\cdots,m)$都是线性函数，则条件(1.2.22)在任何可行点都满足，这时，式(1.2.42)也恒成立. 所以，在这种特殊情形下，K-T点一定是局部极小点. 因为目标函数是线性的，可行域是凸集，所以K-T点也必定是总体极小点.

从而可知，对于一般的非线性规划问题(1.1.1)～(1.1.3)，它的一个K-T点x^*可理解为原问题在x^*处线性化后的一个极小点. 所以，如果需要在K-T点附近找到一个比x^*更好的点(即具有较小目标函数值的可行点)，我们就必须依赖$f(x)$及$C_i(x)$($i=1,\ \cdots,\ m$)的二阶导数. 但是，绝大多数求解问题(1.1.1)～(1.1.3)的数值计算方法均不求二阶导数. 所以，对于这些方法找到了K-T点就达到目的了.

§1.3 方法概述

求解优化问题(1.1.1)～(1.1.3)的计算方法都是迭代法. 即给定一个初始值$x_1\in \mathbf{R}^n$，然后由算法产生迭代点$x_k(k=2,\ 3,\ \cdots)$，希望某一个x_k是K-T点或者$x_k(k=1,\ 2,\ \cdots)$收敛于某一个K-T点.

只有一个变量的优化问题的计算方法是优化计算方法中最基

本的．由于它的特殊性，我们将在第二章对其单独讨论．

对于多维优化问题，现有的计算方法可分为两大类：一类是线搜索(Line search)方法；另一类是信赖域(trust region)方法．

线搜索方法是最常见的、也是研究最多的一类方法．它的基本思想是：在每次迭代中，算法产生一个搜索方向 d_k，一般说来搜索方向 d_k 可看成是某个近似问题的解．我们称该近似问题为子问题，它是原问题在当前迭代点 x_k 附近的一种逼近．算法然后在方向 d_k 上找一个"好"点，也就是寻求一个步长 $\alpha_k>0$，使得点 $x_k+\alpha_k d_k$ 在一定意义下比当前迭代点好．取下一个迭代点 $x_{k+1}=x_k+\alpha_k d_k$．这样就完成了一次迭代．

线搜索方法的本质在于 d_k 的产生与 α_k 的选取．搜索方向 d_k 的产生依赖于子问题的构造．为了使算法有效，无疑要求子问题具有容易求解和较好地逼近原问题等性质．步长 α_k 的确定取决于某种一维搜索技巧和某一评价函数(merit function)．我们认为评价函数的值愈小的点愈"好"．在无约束的情形，显然评价函数可取成目标函数 $f(x)$．在约束优化的情形，评价函数常为一个罚函数(Penalty function)，它依赖于目标函数 $f(x)$ 以及约束函数 $C_i(x)(i=1, 2, \cdots, m)$．

信赖域法是一类较新的方法．它的基本思想是：在每次迭代中，有一个信赖域，它常常是以当前迭代点 x_k 为中心的一个小邻域．在信赖域上求解一个子问题，我们称解 d_k 为试探步(trial step)．然后利用某一评价函数来决定是否接受该试探步以及确定下一次迭代的信赖域．如果试探步被接受，则 $x_{k+1}=x_k+d_k$，否则，$x_{k+1}=x_k$．新的信赖域的大小取决于试探步的好坏．粗略地说，如果试探步较好，则信赖域的大小保持不变或者扩大；否则，将缩小．

为了使子问题便于求解，信赖域常常是以 x_k 为中心、以信赖域半径 $\Delta_k>0$ 为半径的一个广义球 $\{x;\ \|x-x_k\|\leqslant\Delta_k\}$，其中 $\|\cdot\|$ 是 n 维空间的某一范数．$\|\cdot\|$ 一般取为欧氏范数 $\|\cdot\|_2$，$\|\cdot\|_1$ 或者 $\|\cdot\|_\infty$．在这种情形下，调节信赖域的大小等价于调节信赖域半径

的大小.

对信赖域方法的研究主要是从八十年代开始的. 所以它远不如线搜索法那样成熟,应用也没有线搜索法广泛. 但是,由于它有较强的收敛性和可靠性, 近年来对它的研究吸引了不少优化专家和学者. 它必将受到越来越多的人的重视.

优化计算方法还根据算法所需计算目标函数以及约束函数的函数值和导数值分为直接法和非直接法. 直接法是指不需要计算任何导数的方法. 顾名思义,非直接法是指所有不是直接法的方法

直接法由于不需要计算任何导数, 具有简单易用的优点. 因为大量的实际问题中导数相当难求甚至无法求, 所以直接法的应用十分广泛. 但是, 由于直接法不依赖导数, 对它的理论性质进行分析比较困难. 而且在 Powell (1964) 提出共轭方向法后直接法就没有本质性的进展. 随着计算机的飞速发展, 近年来人们对大规模优化问题的兴趣越来越大. 而对于大规模问题, 一般很难求导, 即使能够求导, 工作量也相当大, 所以对直接法的研究将受到更大的重视.

在非直接法中, 除少量方法(如牛顿法)外, 都是仅计算函数值与一阶导数值. 由于非直接法依赖导数, 和直接法相比可得到更多的关于算法的理论性质. 例如一类很重要的非直接法, 拟牛顿法, 在七十年代关于它的收敛性得到了不少优秀的结果. 但是目前关于拟牛顿法的理论分析困难很大, 一是没有新的技巧来进行理论分析, 再就是遗留下来的问题都是难度很大的问题. 另外, 看来想构造出比 BFGS 方法好得多的拟牛顿法似乎是不可能的. 关于非直接法的研究要取得突破性进展, 必须要有新的思想或新的技巧出现才有可能.

§1.4 收敛性与收敛速度

优化计算方法的收敛性与收敛速度是优化方法的重要性质. 给定一个算法和初始点 x_1, 产生的点列 $x_k(k=1, 2, \cdots)$是否收敛

于一个 K-T 点？如果收敛，收敛的速度快不快？这些问题无疑是方法的使用者十分关心的问题．为此，我们给出有关收敛性与收敛速度的一些定义．

设 $x_k\in \mathbf{R}^n(k=1, 2, \cdots)$ 是一无穷点列．令 $\overline{X}$ 是问题(1.1.1)～(1.1.3)的 K-T 点所组成的集合．我们用

$$\operatorname{dist}(x, Y)=\inf_{y\in Y}\|x-Y\|_2 \tag{1.4.1}$$

表示点 x 到集合 Y 的欧氏距离．

定义 1.4.1 如果

$$\liminf_{k\to\infty} \operatorname{dist}(x_k, \overline{X})=0, \tag{1.4.2}$$

我们则称 x_k 子收敛于 $\overline{X}$；如果

$$\lim_{k\to\infty} \operatorname{dist}(x_k, \overline{X})=0, \tag{1.4.3}$$

则称 x_k 弱收敛于 $\overline{X}$；如果存在 $x^*\in\overline{X}$ 使得

$$\lim_{k\to\infty} x_k=x^*, \tag{1.4.4}$$

则称 x_k 强收敛于 $\overline{X}$．

由定义可知，强收敛必是弱收敛，弱收敛必是子收敛．假定 x_k 子收敛于 K-T 点集合 $\overline{X}$，则对任何 $\varepsilon>0$，存在下标 $K(\varepsilon)$ 以及 K-T 点 $x(\varepsilon)\in\overline{X}$，使得 $\|x_{K(\varepsilon)}-x(\varepsilon)\|_2<\varepsilon$．由于在 $x(\varepsilon)$ 处K-T 条件得到满足，故在 $x_{K(\varepsilon)}$ 处 K-T 条件在一定误差范围(与 ε 以及函数性质有关)内得到满足．如果算法所产生的点列 x_k 子收敛于 $\overline{X}$，且在算法中加上终止性判别条件，即当 K-T 条件在一定误差范围内得到满足就停止计算，则算法必定会有限终止．

为了讨论收敛速度，我们假定 $x_k\to x^*$．令 $\varepsilon_k=\|x_k-x^*\|_2$．显然，$x_k$ 收敛于 x^* 的速度就是 ε_k 趋于 0 的速度．假设点列 $x_k(k=1, 2, \cdots)$ 是某一优化计算方法所产生，如果有 k 使得 $x_k=x^*$，则算法必在第 k 次迭代终止．所以我们可假定对于所有的 k 都有 $\varepsilon_k>0$．下列有关收敛速度的定义是由 Ortega 和 Rheinboldt(1970)给出的．

定义 1.4.2 对于 $P\in[1, +\infty)$，我们称

$$Q_p=\limsup_{k\to\infty}\frac{s_{k+1}}{s_k^p} \tag{1.4.5}$$

为点列 x_k 的比值收敛因子，也称 Q-因子；称

$$R_p=\begin{cases}\limsup\limits_{k\to\infty} s_k^{1/k}, & \text{if } p=1;\\ \limsup\limits_{k\to\infty} s_k^{1/p^k}, & \text{if } p>1\end{cases} \tag{1.4.6}$$

为 x_k 的根收敛因子，简称 R- 因子.

定义 1.4.3 我们称

$$O_Q=\inf\{p;\ p\in[1,\ \infty) \text{且} Q_p=+\infty\} \tag{1.4.7}$$

为点列 x_k 的比值收敛阶，简称 Q-收敛阶；称

$$O_R=\inf\{p;\ p\in[1,\ \infty) \text{且} R_p=1\} \tag{1.4.8}$$

为点列 x_k 的根收敛阶，简称 R-收敛阶.

定义 1.4.4 如果 $Q_1=0$，则称 x_k 是 Q-超线性收敛于 x^*；如果 $0<Q_1<1$，则称 x_k 是 Q-线性收敛于 x^*；如果 $Q_1\geqslant1$，则称 x_k 是 Q-次线性收敛于 x^*.

定义 1.4.5 如果 $R_1=0$，则称 x_k 是 R-超线性收敛于 x^*；如果 $0<R_1<1$，则称 x_k 是 R-线性收敛于 x^*；如果 $R_1=1$，则称 x_k 是 R-次线性收敛于 x^*.

定义 1.4.6 如果 $Q_2=0$，则称 x_k 是 Q-超平方收敛于 x^*；如果 $0<Q_2<+\infty$，则称 x_k 是 Q-平方收敛于 x^*；如果 $Q_2=+\infty$，则称 x_k 是 Q-次平方收敛于 x^*.

定义 1.4.7 如果 $R_2=0$，则称 x_k 是 R-超平方收敛于 x^*；如果 $0<R_2<1$，则称 x_k 是 R-平方收敛于 x^*；如果 $R_2\geqslant1$，则称 x_k 是 R-次平方收敛于 x^*.

与定义 1.4.6 和 1.4.7 相似，可定义 R- 和 Q- 立方收敛. 从定义可知，x_k Q-超线性收敛于 x^* 等价于条件

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x^*\|_2}{\|x_k-x^*\|_2}=0. \tag{1.4.9}$$

而(1.4.9)与

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x^*\|_2}{\|x_k-x_{k+1}^*\|_2}=0 \tag{1.4.10}$$

是等价的. 从(1.4.10)式可知，对于 Q-超线性收敛点列 x_k($k=1, 2, \cdots$)，如果 $\|x_k-x_{k+1}\|_2$ 已相当小时，则知 $\|x_{k+1}-x^*\|$ 是一个高阶无穷小.

利用 Q-收敛阶和 R-收敛阶的定义，可以证明以下引理.

引理 1.4.8 对于一收敛点列 x_k，它的 Q-收敛阶不大于其 R-收敛阶，即

$$O_Q \leqslant O_R. \tag{1.4.11}$$

该引理的证明可见 Ortega 和 Rheinboldt(1970).

在以后各章，除特别说明外，我们提到的线性收敛、超线性收敛以及平方收敛分别是指 Q-线性收敛、Q-超线性收敛和 Q-平方收敛.

第 2 章

一维优化方法

一维优化是求解一个单变量函数的极小值，即

$$\min_{x\in\mathbf{R}} f(x), \tag{2.0.1}$$

其中 $f(x)$ 是定义在实数上的非线性函数。一维优化问题是最简单的非线性规划问题，故一维优化方法是非线性规划方法中最基本的，也是最基础的。由于大多数多维优化方法的每次迭代需要进行一维搜索，即精确地或非精确地求解一个多维函数在某个一维子空间上的极小值，而且一维优化的一些方法和技巧可直接推广到多维优化，所以对简单的单变量问题(2.0.1)的算法的研究是有着重要意义的。

假定 $f(x)$ 连续可微，由引理 1.2.1 知问题(2.0.1)的解必为稳定点，即

$$f'(x)=0. \tag{2.0.2}$$

在求解一维优化问题的方法中，有些是基于直接求解(2.0.1)，而有些方法的思想却是基于求(2.0.2)的根。

§2.1 牛 顿 法

牛顿法是一个基于求解(2.0.2)的方法，它之所以被称为牛顿法是因为这一方法可溯源到牛顿提出的求多项式根的一个方法。牛顿方法的思想可从下例中看到：

假定我们要计算 $\sqrt{2}$，它是多项式

$$x^2=2 \tag{2.1.1}$$

的根。设 $\sqrt{2}=1+\varepsilon$，代入上式得

$$1+2\varepsilon+\varepsilon^2=2. \tag{2.1.2}$$

由于 $|\varepsilon|<1$, 我们舍掉上式中的 ε^2 项, 即得

$$\varepsilon\approx 0.5. \tag{2.1.3}$$

从而得到 $\sqrt{2}\approx 1+0.5=1.5$. 如果我们需要更高精度的 $\sqrt{2}$, 令 $\sqrt{2}=1.5+\varepsilon$, 同上可得

$$2.25+3\varepsilon+\varepsilon^2=2. \tag{2.1.4}$$

所以 $\varepsilon\approx -0.25/3$. 从而 $\sqrt{2}\approx 1.5-0.25/3=1.41\dot{6}$. 重复这一步骤, 可求得任意精度的 $\sqrt{2}$ 的近似值.

不难看出, 对于一般的多项式 $P(x)$, 如果 x_k 是 $P(x)=0$ 的一个近似根, 牛顿法的思想是设根为 $x_k+\varepsilon$, 从而可得

$$P(x_k+\varepsilon)=P(x_k)+P'(x_k)\varepsilon+O(\varepsilon^2)=0. \tag{2.1.5}$$

舍去高阶误差项 $O(\varepsilon^2)$, 即知 $\varepsilon\approx -P(x_k)/P'(x_k)$. 故新的近似根为

$$x_{k+1}=x_k-P(x_k)/P'(x_k). \tag{2.1.6}$$

求解多项式根的牛顿方法可推广到求一般非线性方程的根. 我们将其应用到问题(2.0.2)就得到了解一维优化问题的牛顿法.

算法 2.1.1

步 1 初值 x_1 给定, $k=1$.

步 2 计算 $f'(x_k)$, $f''(x_k)$.

步 3 如果 $f'(x_k)=0$, 则停止计算.

$$x_{k+1}=x_k-f'(x_k)/f''(x_k), \tag{2.1.7}$$

$k:=k+1$, 转步 2.

牛顿法的一个优点是它的收敛速度快.

定理 2.1.2 设 $f(x)\in C^2$, $f'(x^*)=0$, $f''(x^*)\neq 0$, 则当 x_1 充分靠近 x^* 时有 $x_k\to x^*$ 且

$$|x_{k+1}-x^*|/|x_k-x^*|\to 0. \tag{2.1.8}$$

如果 $f''(x)$ 是 Hölder 连续, 即存在 $\varepsilon>0$, 使得

$$|f''(x)-f''(y)|\leqslant M|x-y|^\varepsilon. \tag{2.1.9}$$

则

$$|x_{k+1}-x^*|/|x_k-x^*|^{1+\varepsilon}\leqslant \bar{M}. \tag{2.1.10}$$

特别地，如果 $f(x)\in C^3$，则

$$|x_{k+1}-x^*|=O(|x_k-x^*|^2). \tag{2.1.11}$$

证明 利用式(2.1.7)以及 $f'(x^*)=0$，我们可得

$$\begin{aligned}x_{k+1}-x^*&=x_k-x^*-f'(x_k)/f''(x_k)\\&=x_k-x^*-\left[f'(x^*)+\int_{x^*}^{x_k}f''(x)dx\right]/f''(x_k)\\&=\int_{x^*}^{x_k}[f''(x_k)-f''(x)]dx/f''(x_k).\end{aligned} \tag{2.1.12}$$

因为 $f(x)\in C^2$ 以及 $f''(x^*)\neq 0$，必存在 $\delta>0$，使得

$$|f''(x)-f''(y)|<\frac{1}{2}|f''(x^*)| \tag{2.1.13}$$

对任何 x、$y\in[x^*-\delta,\ x^*+\delta]$ 都成立．于是当 $|x_k-x^*|\leqslant\delta$ 时，必有

$$|x_{k+1}-x^*|<\frac{1}{2}|x_k-x^*|\leqslant\delta. \tag{2.1.14}$$

从而可知，当 $|x_1-x^*|\leqslant\delta$ 时，有 $|x_k-x^*|\leqslant\left(\frac{1}{2}\right)^{k-1}\delta$，故 $x_k\to x^*$．而且利用式(2.1.12)、$f''(x^*)\neq 0$ 以及关于 $f(x)$ 的其他假定可证明式(2.1.8)、(2.1.10)和(2.1.11)． ▌

对于收敛序列 $x_k\to x^*$，如果它是线性收敛，根据定义不难证明，必存在一小于 1 的正常数 M，使得

$$|x_{k+1}-x^*|\leqslant M|x_k-x^*|. \tag{2.1.15}$$

对一切充分大的 k 均成立．如果式(2.1.8)成立且存在常数 $\tau>1$，使得

$$|x_{k+1}-x^*|=O(|x_k-x^*|^\tau), \tag{2.1.16}$$

则不难看出 x_k 的 Q-收敛阶不小于 τ．

从定理 2.1.2 可知，只要 $f(x)$ 是三次连续可微的，牛顿法是二次收敛的．对于凸函数 $f(x)$，我们不难证明下面的全局收敛性定理．

定理 2.1.3 设 $f(x)\in C^3$，$f''(x)>0$，且 $f'''(x)$ 不变号，则对任何初始值 x_1，算法 2.1.1 产生的点列 x_k 必有限终止于 $f(x)$ 唯一的极小点 x^*（即对某一 k，$x_k=x^*$）或者二次收敛于 x^*．

对于一般的非线性函数 $f(x)$，算法 2.1.1 仅是一个局部算法，即要求初始点充分靠近某一稳定点．给定同样的初始点，该算法应用到 $\min f(x)$ 和 $\min -f(x)$ 将产生相同的点列，从而 x_k 可能收敛于 $f(x)$ 的极大值点．下面，我们给出一个全局的牛顿算法．

算法 2.1.4

步 1　初始点 x_1 给定，$\delta>0$ 给定，$k=1$．

步 2　计算 $f'(x_k)$，$f''(x_k)$；

　　如果 $f'(x_k)\neq 0$，则转步 4；

　　如果 $f''(x_k)\geqslant 0$，则停；

　　令 $\hat{\delta}:=\delta$．

步 3　$\hat{\delta}:=\hat{\delta}/2$；

　　如果 $f(x_k+\hat{\delta})\geqslant f(x_k)$，则转步 3；

　　$x_{k+1}=x_k+\hat{\delta}$；

　　$k:=k+1$，转步 2．

步 4　$\beta_k=f''(x_k)$，如果 $\beta_k\leqslant 0$，则 $\beta_k=1$；

　　$\alpha_k=1$．

步 5　如果

$$f(x_k-\alpha_k f'(x_k)/\beta_k)\leqslant f(x_k)-\frac{\alpha_k}{4}[f'(x_k)]^2/\beta_k, \tag{2.1.17}$$

　　则转步 6；

　　$\alpha_k:=\alpha_k/2$，转步 5．

步 6　$x_{k+1}=x_k-\alpha_k f'(x_k)/\beta_k$；

　　$k:=k+1$，转步 2．

如果算法进入步 3，则必有 $f'(x_k)=0$，$f''(x_k)<0$，从而对充分小的 $\hat{\delta}>0$，必定有 $f(x_k+\delta)<f(x_k)$，故知步 3 是有限终止的．在步 5 中，由于 $f'(x_k)\neq 0$，利用 Taylor 展开可证明(2.1.17)在 α_k 充分小时一定成立，所以在步 5 中也不会出现无穷循环．

为了证明算法 2.1.4 的全局收敛性，下面给出一个简单的引理：

引理 2.1.5 设 $f(x)\in C^2$, 且对任何 x 有 $|f''(x)|\leqslant M$. 如果存在 $x\in\mathbf{R}$, $\alpha>0$, $c_1\in(0, 1)$, 使得

$$f(x-\alpha f'(x))>f(x)-c_1\alpha[f'(x)]^2. \tag{2.1.18}$$

则必有

$$\alpha>2(1-c_1)/M. \tag{2.1.19}$$

证明 只需利用估计式

$$f(x-\alpha f'(x))\leqslant f(x)-\alpha[f'(x)]^2+\frac{M}{2}\alpha^2[f'(x)]^2 \tag{2.1.20}$$

以及(2.1.18)即可得到结论(2.1.19). ▮

定理 2.1.6 设 $f(x)\in C^2$, 算法 2.1.4 产生的点列 x_k 有界; 如果算法有限终止于 x_k, 则 $f'(x_k)=0$, $f''(x_k)\geqslant 0$, 否则, 对 x_k 的任一聚点 x^*, 均有 $f'(x^*)=0$, $f''(x^*)\geqslant 0$.

证明 显而易见, 如果算法在第 k 次迭代终止, 则 $f'(x_k)=0$, $f''(x_k)\geqslant 0$.

我们假定 x_k 是算法 2.1.4 产生的无穷点列且 $x_k\in[a, b]$ $(k=1, 2, \cdots)$. 令 $M=\max\limits_{x\in[a,b]}|f''(x)|+1$. 首先, 我们有

$$\alpha_k/\beta_k\geqslant\frac{3}{4M}\quad(\forall k). \tag{2.1.21}$$

如果 $\alpha_k=1$, 由于 $\beta_k\leqslant M$, 所以(2.1.21)成立.
若 $\alpha_k<1$, 则不等式(2.1.17)将 α_k 换成 $2\alpha_k$ 后不成立. 利用引理 2.1.5, 可得

$$2\alpha_k/\beta_k\geqslant 2\left(1-\frac{1}{4}\right)\Big/M. \tag{2.1.22}$$

从而可知(2.1.21)恒成立. 所以, 利用(2.1.17)和(2.1.21)可证明, 如果 x_{k+1} 是由步 6 给出, 则必有

$$f(x_{k+1})\leqslant f(x_k)-\frac{3}{16M}[f'(x_k)]^2. \tag{2.1.23}$$

因此可得

$$[f'(x_k)]^2\leqslant\frac{16}{3}M[f(x_k)-f(x_{k+1})]. \tag{2.1.24}$$

如果 x_{k+1} 不是由步 6 给出, 则有 $f'(x_k)=0$. 于是(2.1.24)恒成

立. 因为 $x_k(k=1, 2, \cdots)$ 有界, 且 $f(x_k)$ 单调下降, 故

$$\sum_{k=1}^{\infty}[f'(x_k)]^2\leqslant\sum_{k=1}^{\infty}\frac{16}{3}M[f(x_k)-f(x_{k+1})]<\infty. \tag{2.1.25}$$

所以

$$\lim_{k\to\infty}f'(x_k)=0. \tag{2.1.26}$$

设 x^* 是 x_k 的任一子列 x_{k_i} 的极限点, 由(2.1.26)可知

$$f'(x^*)=0. \tag{2.1.27}$$

不失一般性, 可假定

$$|x_{k_{i+1}}-x^*|<|x_{k_i}-x^*|, \ \forall i. \tag{2.1.28}$$

由(2.1.27)和(2.1.28)以及 $f(x_{k_{i+1}})<f(x_{k_i})$ 可证

$$f''(x^*)\geqslant 0. \tag{2.1.29}$$

从而可知定理成立. ▌

算法 2.1.4 要求 $f'(x_k)=0$ 才停止计算, 否则, 将无穷迭代下去. 在具体计算时, 给定一个容许误差 $\varepsilon>0$, 当 $|f'(x_k)|\leqslant\varepsilon$ 时, 即可停机.

§2.2 割 线 法

牛顿法需要计算二阶导数. 利用差商代替导数, 即

$$f''(x_k)\approx\frac{f'(x_k)-f'(x_{k-1})}{x_k-x_{k-1}} \tag{2.2.1}$$

从牛顿法可导出下面的迭代公式:

$$x_{k+1}=x_k-\frac{f'(x_k)(x_k-x_{k-1})}{f'(x_k)-f'(x_{k-1})}. \tag{2.2.2}$$

割线法的基本迭代就是用(2.2.2)式, 它在几何上是求通过 $(x_{k-1}, f'(x_{k-1}))$、$(x_k, f'(x_k))$ 两点的曲线 $y=f'(x)$ 的割线的零点. 假定 $f(x)\in C^3$, x^* 是 $f(x)$ 的稳定点且 $f''(x^*)\neq 0$. 在 x^* 附近, 对割线法有如下估计:

$$x_{k+1}-x^*=x_k-x^*-\frac{f'(x_k)(x_k-x_{k-1})}{f'(x_k)-f'(x_{k-1})}$$

$$=\frac{1}{f'(x_k)-f'(x_{k-1})}[f'(x_k)(x_{k-1}-x^*)-f'(x_{k-1})(x_k-x^*)]$$

$$=\frac{x_k-x_{k-1}}{f'(x_k)-f'(x_{k-1})}\Big[\frac{f'''(x^*)}{2}(x_k-x^*)(x_{k-1}-x^*)$$

$$+O((x_k-x^*)^2|x_{k-1}-x^*|+(x_{k-1}-x^*)^2|x_k-x^*|)\Big]$$

$$=\frac{f'''(x^*)}{2f''(x^*)}(x_k-x^*)(x_{k-1}-x^*)+o(|x_k-x^*||x_{k-1}-x^*|).$$

(2.2.3)

利用上式，我们可证存在 $\delta>0$，当 x_1、$x_2\in(x^*-\delta,\ x^*+\delta)$ 且 $x_1\neq x_2$ 时，$x_k\to x^*$．而且

$$|x_{k+1}-x^*|=\beta_k|x_k-x^*||x_{k-1}-x^*|, \tag{2.2.4}$$

其中

$$\beta_k=\left|\frac{f'''(x^*)}{2f''(x^*)}\right|+o(1)。 \tag{2.2.5}$$

为了导出割线法的收敛速度，我们给出如下引理：

引理 2.2.1 设 $\varepsilon_k>0(k=1,\ 2,\ \cdots)$是一趋于 0 的数列且满足

$$\varepsilon_{k+1}=\beta_k\varepsilon_k\varepsilon_{k-1}, \tag{2.2.6}$$

其中

$$\beta_k\to\beta^*>0, \tag{2.2.7}$$

则必有

$$\lim_{k\to\infty}\varepsilon_{k+1}/\varepsilon_k^{\tau}=(\beta^*)^{1/\tau}, \tag{2.2.8}$$

$\tau=(\sqrt{5}+1)/2$ 是方程 $t^2=t+1$ 的根．

证明 从(2.2.7)可得

$$\lg\varepsilon_{k+1}=\lg\varepsilon_k+\lg\varepsilon_{k-1}+\lg\beta_k. \tag{2.2.9}$$

因为 $\tau^2=\tau+1$，故

$$\lg\varepsilon_{k+1}-\tau\lg\varepsilon_k=-\frac{1}{\tau}[\lg\varepsilon_k-\tau\lg\varepsilon_{k-1}]+\lg\beta_k.$$

(2.2.10)

定义 $\eta_k=\lg\varepsilon_{k+1}-\tau\lg\varepsilon_k$．由于 $\beta_k\to\beta^*$，$\tau>1$，从(2.2.10)式可

知 $\eta_k(k=1, 2, \cdots)$ 必是一有界数列，且满足

$$\eta_{k+1}-\eta_k=-\frac{1}{\tau}(\eta_k-\eta_{k-1})+\lg\frac{\beta_{k+1}}{\beta_k}. \tag{2.2.11}$$

于是 $\eta_{k+1}-\eta_k\to 0$. 假定 $\eta_{k_i}\to\eta^*$ 是 η_k 的任一收敛子列，因为 $\eta_{k+1}-\eta_k\to 0$，必有 $\eta_{k_i+1}\to\eta^*$. 因而从(2.2.10)式可得

$$\eta^*=-\frac{1}{\tau}\eta^*+\lg\beta^*, \tag{2.2.12}$$

所以

$$\eta^*=\frac{1}{\tau}\lg\beta^*. \tag{2.2.13}$$

由于 η_{k_i} 是任一收敛子列，故知 η_k 收敛，且

$$\lim_{k\to\infty}\eta_k=\frac{1}{\tau}\lg\beta^*. \tag{2.2.14}$$

注意到 η_k 的定义，有

$$\lim_{k\to\infty}\varepsilon_{k+1}/\varepsilon_k^{\tau}=\lim_{k\to\infty}e^{\eta_k}=(\beta^*)^{\frac{1}{\tau}}. \tag{2.2.15}$$

所以引理成立. ∎

从引理 2.2.1 以及式(2.2.4)和式(2.2.5)可知，如果 $f'''(x^*)\neq 0$，且当 $x_1\neq x_2$ 充分靠近 x^* 时，割线法产生的点列 x_k 必超线性收敛于 x^*，且

$$\lim_{k\to\infty}\frac{|x_{k+1}-x^*|}{|x_k-x^*|^{\tau}}=\left|\frac{f'''(x^*)}{2f''(x^*)}\right|^{1/\tau}. \tag{2.2.16}$$

从上述可知，割线法是 τ 阶收敛于 x^* 的，它显然比牛顿法收敛得慢. 利用(2.1.12)式，可证明，对于牛顿法，极限

$$\lim_{k\to\infty}\frac{|x_{k+1}-x^*|}{|x_k-x^*|^2}=\left|\frac{f'''(x^*)}{2f''(x^*)}\right|. \tag{2.2.17}$$

成立. 所以当 k 充分大时，对割线法有如下估计：

$$|x_{k+N}-x^*|\approx\left|\frac{f'''(x^*)}{2f''(x^*)}\right|^{\frac{1}{\tau}+\frac{(N-1)N}{2}}|x_k-x^*|^{\tau N}, \tag{2.2.18}$$

而对牛顿法，有

$$|x_{k+N}-x^*|\approx\left|\frac{f'''(x^*)}{2f''(x^*)}\right|^{1+N(N-1)}|x_k-x^*|^{2N}. \tag{2.2.19}$$

所以，当 k 充分大时，要使 x_k-x^* 减小一定量级割线法所需迭代次数 N_1 与牛顿法所需迭代次数 N_2 之比 $N_1/N_2\approx 2/\tau$. 但是割线法每次迭代只需求一个值 $f'(x_k)$，而牛顿法需要计算 $f'(x_k)$、$f''(x_k)$. 所以，牛顿法所需的求值次数约是割线法的 τ 倍. 下面看一个例子：

$$\min_{x\in\mathbf{R}} -xe^{-x} \tag{2.2.20}$$

问题(2.2.20)有唯一的极小点 $x^*=1$. 取初始值 $x_1=0$，用牛顿法(2.1.7)可得 $x_2=0.5$. 所以，用割线法时，取 $x_1=0$, $x_2=0.5$. 下面表 2.2.1 给出牛顿法(2.1.7)和割线法(2.2.2)的迭代过程. 为了便于比较收敛速度，在表中只给出解与迭代点之差 $1-x_k$. 迭代终止的条件是：

$$|f'(x_k)|\leqslant 0.5\times 10^{-11}. \tag{2.2.21}$$

表 2.2.1　牛顿法与割线法的比较：$f(x)=-xe^{-x}$

	(2.1.7)	(2.2.2)
$1-x_1$	1.0	1.0
$1-x_2$	0.5	0.5
$1-x_3$	0.16666666	0.28236670
$1-x_4$	0.23809524×10^{-1}	0.10119747
$1-x_5$	0.55370986×10^{-3}	0.23921656×10^{-1}
$1-x_6$	0.30642493×10^{-6}	0.22772164×10^{-2}
$1-x_7$	$0.93897112\times 10^{-13}$	0.53768353×10^{-4}
$1-x_8$		0.12229958×10^{-6}
$1-x_9$		$0.65756706\times 10^{-11}$

从表 2.2.1 可知，牛顿法只需 6 次迭代，而割线法需要 8 次迭代. 牛顿法计算函数导数、二阶导数的次数与割线法计算函数导数的次数之比为 $(2\times 6)/8=1.5$，这与我们所估计的比值 τ 相差不大.

牛顿法(2.1.7)产生的 x_{k+1} 是近似函数

$$\hat{f}_k(x)=f(x_k)+f'(x_k)(x-x_k)+\frac{1}{2}f''(x_k)(x-x_k)^2\approx f(x) \tag{2.2.22}$$

的稳定点. 当 $f''(x_k)>0$ 时，x_{k+1} 是原目标函数 $f(x)$ 在 x_k 点的二

阶泰勒展开的极小点．下面考虑一般形式的二次逼近函数

$$\varphi_k(x, c_k)=f(x_k)+f'(x_k)(x-x_k)+\frac{1}{2}c_k(x-x_k)^2, \tag{2.2.23}$$

其中 c_k 是 $f''(x_k)$ 的近似值．令 x_{k+1} 是(2.2.23)的稳定点，即得

$$x_{k+1}=x_k-f'(x_k)/c_k. \tag{2.2.24}$$

$\varphi_k(x, c_k)$是 $f(x)$在 x_k 点附近的一个逼近函数，且满足插值条件

$$\varphi_k(x_k, c_k)=f(x_k) \tag{2.2.25}$$

$$\varphi_k'(x_k, c_k)=f'(x_k). \tag{2.2.26}$$

如果要求

$$\varphi_k'(x_{k-1}, c_k)=f'(x_{k-1}), \tag{2.2.27}$$

则 $c_k=[f'(x_k)-f'(x_{k-1})]/(x_k-x_{k-1})$，这时(2.2.24)即是割线法(2.2.2)．

我们把(2.2.27)换成

$$\varphi_k(x_{k-1}, c_k)=f(x_{k-1}), \tag{2.2.28}$$

则 $c_k=2[f'(x_k)-(f(x_k)-f(x_{k-1}))/(x_k-x_{k-1})]/(x_k-x_{k-1})$，从而得到迭代法：

$$x_{k+1}=x_k-\frac{f'(x_k)[x_k-x_{k-1}]}{2[f'(x_k)-(f(x_k)-f(x_{k-1}))/(x_k-x_{k-1})]} \tag{2.2.29}$$

与(2.2.3)相似，我们可证

$$x_{k+1}-x^*=\frac{f'''(x^*)}{3f''(x^*)}(x_k-x^*)(x_{k-1}-x^*)+o(|x_k-x^*||x_{k-1}-x^*|). \tag{2.2.30}$$

所以，设 $f'(x^*)=0$, $f''(x^*)\neq 0$, $f'''(x^*)\neq 0$，当 x_1、x_2 充分靠近 x^* 时有

$$\lim_{k\to\infty}|x_{k+1}-x^*|/|x_k-x^*|^\tau=\left|\frac{f'''(x^*)}{3f''(x^*)}\right|^{1/\tau}. \tag{2.2.31}$$

其中 $\tau=(\sqrt{5}+1)/2$．比较(2.2.17)和(2.2.31)即知迭代法(2.2.29)比割线法(2.2.2)稍快．

一般说来，二次函数(2.2.23)不可能同时满足四个插值条件

(2.2.25)~(2.2.28). 我们考虑三次 Hermit 插值函数:

$$\begin{aligned}\tilde{\varphi}_k(x)=&f(x_{k-1})(x-x_k)^2(2x+x_k-3x_{k-1})/(x_k-x_{k-1})^3\\&+f(x_k)(x-x_{k-1})^2(3x_k-2x-x_{k-1})/(x_k-x_{k-1})^3\\&+f'(x_{k-1})(x-x_k)^2(x-x_{k-1})/(x_k-x_{k-1})^2\\&+f'(x_k)(x-x_{k-1})^2(x-x_k)/(x_k-x_{k-1})^2.\end{aligned}\tag{2.2.32}$$

通过直接计算可得

$$\tilde{\varphi}''_k(x_k)=\left[4f'(x_k)+2f'(x_{k-1})-6\,\frac{f(x_k)-f(x_{k-1})}{x_k-x_{k-1}}\right]\Big/(x_k-x_{k-1}).\tag{2.2.33}$$

我们用 $\tilde{\varphi}''_k(x_k)$代替牛顿法中的$f''(x_k)$就得到下面的迭代公式

$$x_{k+1}=x_k-\frac{f'(x_k)(x_k-x_{k-1})}{4f'(x_k)+2f'(x_{k-1})-6[f(x_k)-f(x_{k-1})]/(x_k-x_{k-1})}.\tag{2.2.34}$$

设$f'(x^*)=0$, $f''(x^*)\neq0$, x_k、x_{k-1}充分靠近x^*,我们可证

$$\begin{aligned}x_{k+1}-x^*=&\frac{f'''(x^*)}{2f''(x^*)}(x_k-x^*)^2-\frac{f^{(4)}(x^*)}{12f''(x^*)}(x_k-x^*)(x_k-x_{k-1})^2\\&+O(|x_k-x^*|^3+|x_k-x^*|^2|x_k-x_{k-1}|^2\\&+|x_k-x^*||x_k-x_{k-1}|^3).\end{aligned}\tag{2.2.35}$$

从上式可知迭代法(2.2.34)在x^*附近超线性收敛. 于是,有

$$x_{k+1}-x^*=\alpha_k(x_k-x^*)^2+\beta_k(x_k-x^*)(x_{k-1}-x^*)^2.\tag{2.2.36}$$

其中

$$\alpha_k=\frac{f'''(x^*)}{2f''(x^*)}+O(|x_k-x^*|),\tag{2.2.37}$$

$$\beta_k=-\frac{f^{(4)}(x^*)}{12f''(x^*)}+O(|x_{k-1}-x^*|).\tag{2.2.38}$$

为了分析(2.2.36)的收敛速度,我们给出如下一般性结果:

引理 2.2.2 设$e_k>0(k=1,2,\cdots)$是一趋于0的数列且满足关系式

$$e_{k+1}=\bar{\beta}_k e_k e_{k-1}^2,\quad\forall k>1,\tag{2.2.39}$$

其中$\bar{\beta}_k\to\bar{\beta}^*>0$,而且

$$\sum_{k=2}^{\infty}|\bar{\beta}_k-\bar{\beta}^*|<\infty,\tag{2.2.40}$$

则必存在 $\overline{M}_1>\overline{M}_2>0$, 使得对一切 k 都有

$$\overline{M}_2\leqslant\varepsilon_{k+1}/\varepsilon_k^2<\overline{M}_1. \tag{2.2.41}$$

证明 令 $\eta_k=\varepsilon_{k+1}/\varepsilon_k^2$, 由(2.2.39)式可知

$$\eta_{k+2}=\overline{\beta}_{k+1}/\eta_{k+1}=(\overline{\beta}_{k+1}/\overline{\beta}_k)\eta_k. \tag{2.2.42}$$

于是

$$\eta_{2k+1}=\prod_{i=1}^{k}(\overline{\beta}_{2i+1}/\overline{\beta}_{2i})\eta_1. \tag{2.2.43}$$

因为级数 $\sum\limits_{k=2}^{\infty}|\overline{\beta}_k-\overline{\beta}^*|<\infty$, 所以无穷乘积 $\prod\limits_{i=1}^{\infty}(\overline{\beta}_{2i+1}/\overline{\beta}_{2i})$ 也必收敛. 所以

$$\lim_{k\to\infty}\eta_{2k+1}=\eta^*>0. \tag{2.2.44}$$

利用上式与(2.2.42)式可推出

$$\lim_{k\to\infty}\eta_{2k}=\overline{\beta}^*/\eta^*>0. \tag{2.2.45}$$

从(2.2.44)式和(2.2.45)式即知必存在 $M_1>M_2>0$ 使得(2.2.41)式成立. ▮

利用引理 2.2.2, 我们可证下列收敛性定理:

定理 2.2.3 设迭代法(2.2.34)产生的点列 $x_k\to x^*$, $f'(x^*)=0$, $f''(x^*)\neq0$, $f'''(x^*)=0$, $f^{(4)}(x^*)\neq0$, 则存在 $M_1>M_2>0$, 使得

$$M_2|x_{k+1}-x^*|^2\leqslant|x_{k+1}-x^*|\leqslant M_1|x_k-x^*|^2. \tag{2.2.46}$$

证明 因为 $(x_{k+1}-x^*)=o(|x_k-x^*|)$, 从(2.2.36)以及假设条件可知

$$|x_{k+1}-x^*|=\hat{\beta}_k|x_k-x^*|\,|x_{k-1}-x^*|^2, \tag{2.2.47}$$

其中

$$\hat{\beta}_k=\left|\frac{f^{(4)}(x^*)}{12f''(x^*)}\right|+O(|x_{k-1}-x^*|). \tag{2.2.48}$$

由于 x_k 超线性收敛, 所以 $\sum\limits_{k=1}^{\infty}|x_k-x^*|<\infty$. 因而从引理 2.2.2 和(2.2.47)、(2.2.48)即知定理成立. ▮

定理 2.2.4 设迭代法(2.2.34)产生的点列 $x_k\to x^*$, $f'(x^*)=0$, $f''(x^*)\neq0$, $f'''(x^*)\neq0$, $f^{(4)}(x^*)=0$, 则必有

$$\limsup_{k\to\infty}\frac{|x_{k+1}-x^*|}{|x_k-x^*|^2}\leqslant\left|\frac{f'''(x^*)}{2f''(x^*)}\right|. \tag{2.2.49}$$

证明　定义

$$\eta_k=|x_{k+1}-x^*|/|x_k-x^*|^2,\ \alpha^*=|f'''(x^*)/2f''(x^*)|,$$

由(2.2.35)式及假定 $f^{(4)}(x^*)=0$ 可知

$$\eta_k=\alpha^*+O(|x_{k-1}-x^*|/\eta_{k-1}+|x_k-x^*|). \tag{2.2.50}$$

因为 $x_k\to x^*$ 且 $\alpha^*>0$, 存在 k_0 以及 $\delta>0$ 使得当 $k\geqslant k_0$ 且 $\eta_k\geqslant\alpha^*-\delta>0$ 时, 必有 $|\eta_{k+1}-\alpha^*|<\delta$.

假定(2.2.49)不成立, 则必存在子列 k_i, 使得 $\eta_{k_i}>\alpha^*$. 故存在 j, 使得 $k_j>k_0$ 且 $\eta_{k_j}>\alpha^*$. 由上面的分析, 利用归纳法可证, $|\eta_k-\alpha^*|<\delta(\forall k>k_j)$. 由(2.2.50)即知 $\eta_k\to\alpha^*$, 这与假定(2.2.49)不成立相矛盾, 此矛盾说明了定理正确. ▌

比较(2.2.17)和(2.2.49)即知在 $f^{(4)}(x^*)=0$ 的假定下迭代法(2.2.34)的收敛速度不比牛顿法慢.

定理 2.2.5　设迭代法(2.2.34)产生的点列 $x_k\to x^*$, $f'(x^*)=0$, $f''(x^*)\neq0$, $f'''(x^*)\neq0$, $f^{(4)}(x^*)\neq0$, 则有

$$|x_{k+1}-x^*|=O(|x_k-x^*|^2+|x_{k-1}-x^*|^4). \tag{2.2.51}$$

如果 $f'''(x^*)f^{(4)}(x^*)<0$, 则存在 $M>0$, 使得

$$|x_{k+1}-x^*|\leqslant M|x_k-x^*|^2. \tag{2.2.52}$$

如果 $f'''(x^*)f^{(4)}(x^*)>0$ 且 $|f'''(x^*)|^2>\frac{4}{3}|f''(x^*)||f^{(4)}(x^*)|$, 则也存在 $M>0$ 使得(2.2.52)成立.

证明　令 $\eta_k=(x_{k+1}-x^*)/(x_k-x^*)^2$, 从(2.2.36)可知

$$\eta_k=\alpha_k+\beta_k/\eta_{k-1}. \tag{2.2.53}$$

为了证明简单, 我们不妨假定

$$\eta_k=\alpha^*+\beta^*/\eta_{k-1}, \tag{2.2.54}$$

其中 $\alpha^*=f'''(x^*)/2f''(x^*)$, $\beta^*=-f^{(4)}(x^*)/12f''(x^*)$. 从(2.2.54)式可知如果 $|\eta_k|>\frac{3}{2}|\alpha^*|$, 则必有 $|\eta_{k-1}|<2|\beta^*/\alpha^*|$. 于是有

$$|x_{k+1}-x^*|\leqslant\frac{3}{2}|\alpha^*|\cdot|x_k-x^*|^2 \tag{2.2.55}$$

或者

$$|x_{k+1}-x^*|\leqslant 6(\beta^*)^3/|\alpha^*|\cdot|x_{k-1}-x^*|^4, \tag{2.2.56}$$

于是(2.2.51)式成立.

现在证明定理的其余部分. 由于证明完全相同, 我们仅证明 $\alpha^*>0$ 的情形.

如果 $f'''(x^*)f^{(4)}(x^*)<0$, 这时必有 $\beta^*>0$, 从而函数 $\psi(t)=\alpha^*+\beta^*/t$ 是 $(-\infty, 0)$ 和 $(0, +\infty)$ 上的严格单调下降函数. 如果 $\eta_k<0(\forall k)$, 则知 $\eta_k>-\beta^*/\alpha^*$, 于是(2.2.52)式对 $M=\beta^*/\alpha^*$ 成立; 否则, 必存在 k_0 使得 $\eta_{k_0}>0$. 由(2.2.54)知 $\eta_k>\alpha^*(\forall k\geqslant k_0)$. 于是当 k 大于 k_0 时, $\eta_k<\alpha^*+\beta^*/\alpha^*$, 故知(2.2.52)式成立.

对于 $f'''(x^*)f^{(4)}(x^*)>0$, 这时 $\beta^*<0$, $\psi(t)$ 是严格单调上升函数. 假定 $|f'''(x^*)|^2>\frac{4}{3}|f''(x^*)||f^{(4)}(x^*)|$, 故有 $(\alpha^*)^2+4\beta^*>0$. 令

$$t^*=(\alpha^*+\sqrt{(\alpha^*)^2+4\beta^*})/2, \tag{2.2.57}$$

则知

$$\psi(t^*)=t^*, \tag{2.2.58}$$

且

$$t^*<\psi(t)<t, \ \forall t>t^*. \tag{2.2.59}$$

如果 $\eta_k<t^*(\forall k)$, 则从(2.2.54)可知

$$\eta_k>(t^*-\alpha^*)/\beta^*>0, \tag{2.2.60}$$

从而(2.2.52)成立. 如果存在 k_0, 使得 $\eta_{k_0}\geqslant t^*$, 则由(2.2.58)和(2.2.59)式可得 $\lim\limits_{k\to\infty}\eta_k=t^*$, 所以(2.2.52)式必成立. ▌

利用(2.2.51)式可证明, 存在 $M>0$, $q<1$, 使得对一切 k,

$$|x_k-x^*|\leqslant Mq^{2^k}. \tag{2.2.61}$$

由定义1.4.3即知点列 x_k 的 R-收敛阶不小于2. 由于牛顿法的 R-收敛阶为2, 所以我们可知迭代法(2.2.34)在 R-收敛的意义下不比牛顿法慢.

我们用迭代公式(2.2.34)和迭代公式(2.2.29)分别求解问题(2.2.20), 得到的结果如表2.2.2所列.

表 2.2.2 迭代法(2.2.34)与(2.2.29)的计算表现：$f(x)=-xe^{-x}$.

	(2.2.34)	(2.2.29)
$1-x_1$	1.0	1.0
$1-x_2$	0.5	0.5
$1-x_3$	0.14415722	0.25000000
$1-x_4$	0.13581091×10^{-1}	0.78248732×10^{-1}
$1-x_5$	0.12099140×10^{-3}	0.13320287×10^{-1}
$1-x_6$	0.91183707×10^{-8}	0.72988299×10^{-3}
$1-x_7$	0.11102230×10^{-15}	0.66295735×10^{-5}
$1-x_8$		0.32398203×10^{-8}
$1-x_9$		0.12226831×10^{-13}

从表 2.2.1 和表 2.2.2 可知，迭代法(2.2.34)是二阶收敛的，它的计算表现和牛顿法相似；迭代法(2.2.29)是 τ 阶收敛的，它比割线法(2.2.2)收敛稍快. 这些计算结果和我们的理论分析完全吻合.

§2.3 多项式插值法

在上节讨论的割线法和迭代法(2.2.29)均可看成是利用二次插值推导的. 在本节我们给出另两个常见的多项式插值导出的方法，一个是基于三点函数值的二次插值；另一个是基于两点的 Hermit 插值.

首先，我们考虑在 x_k, x_{k-1}, x_{k-2} 三点上的二次插值函数. 显然，我们可将插值函数写成牛顿形式：

$$\begin{aligned}\hat{f}_k(x)=&f(x_k)+(x-x_k)f(x_k,\ x_{k-1})\\&+(x-x_k)(x-x_{k-1})f(x_k,\ x_{k-1},\ x_{k-2}),\end{aligned}\qquad(2.3.1)$$

其中

$$f(x_k,\ x_{k-1})=[f(x_k)-f(x_{k-1})]/(x_k-x_{k-1}),\qquad(2.3.2)$$

$$f(x_k,\ x_{k-1},\ x_{k-2})=[f(x_k,\ x_{k-1})-f(x_{k-1},\ x_{k-2})]/(x_k-x_{k-2}),\qquad(2.3.3)$$

是函数 $f(x)$ 的一阶及二阶差商. 设 $f(x_k,\ x_{k-1},\ x_{k-2})\neq0$, 令 x_{k+1}

是(2.3.1)的稳定点，即得到迭代法：

$$x_{k+1}=x_k-\frac{f(x_k,\ x_{k-1})+(x_k-x_{k-1})f(x_k,\ x_{k-1},\ x_{k-2})}{2f(x_k,\ x_{k-1},\ x_{k-2})}$$

$$=\frac{1}{2}(x_k+x_{k-1})-\frac{f(x_k,\ x_{k-1})}{2f(x_k,\ x_{k-1},\ x_{k-2})}. \quad (2.3.4)$$

迭代法(2.3.4)是基于二次插值(2.3.1)，故被称为抛物线法.

设 $x_k,\ x_{k-1},\ x_{k-2}$ 充分靠近 x^*，且互不相等，$f'(x^*)=0$，$f''(x^*)\neq 0$. 记 $\varepsilon_i=x_i-x^*(\forall i)$，则从(2.3.4)式可证明

$$\varepsilon_{k+1}=\frac{f'''(x^*)}{6f''(x^*)}[\varepsilon_k\varepsilon_{k-1}+\varepsilon_{k-2}\varepsilon_k+\varepsilon_{k-2}\varepsilon_{k-1}]$$

$$+O(\max[|\varepsilon_k|^3,\ |\varepsilon_{k-1}|^3,\ |\varepsilon_{k-2}|^3]). \quad (2.3.5)$$

由上式可证明 ε_k 是 R-超线性收敛，且它的 R-收敛阶不小于 $\bar{\tau}$，其中 $\bar{\tau}$ 是方程 $t^3=t+1$ 的正根，即

$$\bar{\tau}=[\sqrt[3]{1-\sqrt{23/27}}+\sqrt[3]{1+\sqrt{23/27}}]/\sqrt[3]{2}\approx 1.32. \quad (2.3.6)$$

假定 ε_k 是 Q-超线性收敛，则从(2.3.5)式可得

$$\varepsilon_{k+1}=\frac{f'''(x^*)}{6f''(x^*)}\varepsilon_{k-2}\varepsilon_{k-1}+o(|\varepsilon_{k-2}|\cdot|\varepsilon_{k-1}|). \quad (2.3.7)$$

如果 $f'''(x^*)\neq 0$，与引理 2.2.1 类似，从(2.3.7)式可证明

$$\lim_{k\to\infty}\frac{|\varepsilon_{k+1}|}{|\varepsilon_k^{\bar{\tau}}|}=\left|\frac{f'''(x^*)}{6f''(x^*)}\right|^{1/\bar{\tau}}. \quad (2.3.8)$$

为了使抛物线法实用化，我们要求在每次迭代都有

$$f(x_k,\ x_{k-1},\ x_{k-2})>0. \quad (2.3.9)$$

下面的算法找出了三个初始点 x_1、x_2、x_3 使得 $x_3<x_1<x_2$ 或者 $x_2<x_1<x_3$，而且有

$$f(x_2)>f(x_1),\ f(x_1)>f(x_3). \quad (2.3.10)$$

由于 x_1 在 x_2 与 x_3 之间且(2.3.10)满足，故必有(2.3.9). 因为该算法主要利用划界(bracketing)技巧，所以它被称为划界算法.

算法 2.3.1 (划界算法)

步 1 给定 $x_0\in\mathbf{R}$，$h>0$，$\alpha>1$；$x_2=x_0$，$x_1=x_0+h$.

步 2 如果 $f(x_2)>f(x_1)$，则转步 4；

如果 $f(x_2)<f(x_1)$，则转步 6.

步 3　$x_1 := \frac{1}{2}(x_1+x_2)$，如果$f(x_2)=f(x_1)$，则转步 3；

如果$f(x_1)>f(x_2)$，则转步 6；

$x_3=2x_1-x_2$，停止.

步 4　$h:=2h$，$x_3=x_1+h$，如果$f(x_3)>f(x_1)$，则停；

如果$f(x_3)=f(x_1)$，则转步 5；

$x_2:=x_1$，$x_1:=x_3$，转步 4.

步 5　$\hat{x}:=(x_3+x_1)/2$；如果$f(\hat{x})>f(x_1)$，则$x_3=\hat{x}$及停；

如果$f(\hat{x})<f(x_1)$，则$x_2:=x_1$，$x_1:=\hat{x}$及停；

$x_3:=\hat{x}$，转步 5.

步 6　$\hat{x}=x_1$，$x_1=x_2$，$x_2=\hat{x}$，$h=-h$，转步 4.

不难看出，上面的算法如果不有限终止，则必产生无穷点列$x_3^{(k)}$或者产生无穷点列$\hat{x}^{(k)}$. 在前一种情况，有$|x_3^{(k)}|\to+\infty$且$f(x_3^{(k)})$严格单调下降，此时可理解函数$f(x)$在无穷远处达到极小. 在后一种情况，$\hat{x}^{(k)}$收敛于一有限点$\bar{x}$，且$f(\hat{x}^{(k)})=f(\bar{x})$，于是$\bar{x}$是$f(x)$的稳定点. 当算法有限终止时，$x_1$必介于$x_2$与$x_3$之间且不等式(2.3.10)成立，从而(2.3.9)式成立.

下面的算法基于抛物线法迭代公式(2.3.4). 在每次迭代利用适当选取迭代点使得下一次迭代所需的三个初始点满足(2.3.10).

算法 2.3.2(抛物线法)

步 1　给定$x_0\in\mathbf{R}$，$\varepsilon>0$，由算法 2.3.1 产生x_1，x_2，x_3.

步 2　如果$\min\{|x_2-x_1|, |x_3-x_1|\}<\varepsilon$，则停.

用迭代公式(2.3.4)计算x_4.

步 3　如果$f(x_4)<f(x_1)$，则转步 5.

如果$f(x_4)>f(x_1)$，则转步 6.

步 4　$\hat{x}=\frac{1}{2}(x_1+x_4)$；如果$\hat{x}=x_1$，则$\hat{x}=\frac{1}{2}(x_1+x_2)$；

如果$f(\hat{x})=f(x_1)$，则$x_4=\hat{x}$，转步 4；

$x_4=\hat{x}$，转步 3.

步 5　如果$(x_4-x_1)(x_3-x_1)>0$，则$x_2=x_1$，否则，$x_3=x_1$；

$x_1=x_4$, 转步 2.

步 6 如果$(x_4-x_1)(x_3-x_1)>0$, 则 $x_3=x_4$, 否则, $x_2=x_4$; 转步 2.

下面介绍一个基于三次插值的迭代公式. 考虑基于 x_k、x_{k-1} 两点的三次 Hermit 插值(2.2.32). 假定

$$f'(x_{k-1})(x_k-x_{k-1})<0,\ f'(x_k)(x_{k-1}-x_k)<0. \tag{2.3.11}$$

显然(2.3.11)等价于

$$x_{k-1}<x_k,\ f'(x_{k-1})<0<f'(x_k), \tag{2.3.12}$$

或者

$$x_k<x_{k-1},\ f'(x_k)<0<f'(x_{k-1}). \tag{2.3.13}$$

在(2.3.11)的假定下, 函数$f(x)$在 x_{k-1} 与 x_k 之间必有局部极小点. 而且插值函数(2.2.32)在 x_{k-1} 与 x_k 之间存在唯一的稳定点 x_{k+1}, 该稳定点是一个局部严格极小点. 因为 $\tilde{\varphi}'_k(x_{k+1})=0$, 从而有

$$\begin{aligned}
&6(x_{k+1}-x_k)(x_{k+1}-x_{k-1})[f(x_{k-1})\\
&-f(x_k)]/(x_k-x_{k-1})^3+f'(x_{k-1})[(x_{k+1}-x_k)^2\\
&+2(x_{k+1}-x_k)(x_{k+1}-x_{k-1})]/(x_k-x_{k-1})^2\\
&+f'(x_k)[(x_{k+1}-x_{k-1})^2\\
&+2(x_{k+1}-x_{k-1})(x_{k+1}-x_{k-1})]/(x_k-x_{k-1})^2=0.
\end{aligned} \tag{2.3.14}$$

令 $t_k=(x_{k+1}-x_k)/(x_{k-1}-x_k)$, 则得到

$$3t_k^2\alpha_k-2t_k\beta_k+\gamma_k=0. \tag{2.3.15}$$

其中

$$\alpha_k=f'(x_k)+f'(x_{k-1})-2f(x_k,\ x_{k-1}), \tag{2.3.16}$$

$$\beta_k=2f'(x_k)+f'(x_{k-1})-3f(x_k,\ x_{k-1}), \tag{2.3.17}$$

$$\gamma_k=f'(x_k). \tag{2.3.18}$$

如果 $\alpha_k=0$, 则有

$$x_{k+1}=x_k-\frac{\gamma_k(x_k-x_{k-1})}{2\beta_k}. \tag{2.3.19}$$

这就是在上节中讨论过的(2.2.34). 现假定 $\alpha_k\neq0$, 由于 $0<t_k<1$, 从(2.3.15)式可知

$$\begin{aligned}
t_k&=(\beta_k+\sqrt{\beta_k^2-3\alpha_k\gamma_k})/3\alpha_k\\
&=-\gamma_k/(\sqrt{\beta_k^2-3\alpha_k\gamma_k}-\beta_k)\\
&=\frac{-\gamma_k(\sqrt{\beta_k^2-3\alpha_k\gamma_k}+\beta_k-2\gamma_k)}{(\sqrt{\beta_k^2-3\alpha_k\gamma_k}-\beta_k)(\sqrt{\beta_k^2-3\alpha_k\gamma_k}+\beta_k-2\gamma_k)}\\
&=\frac{\sqrt{\beta_k^2-3\alpha_k\gamma_k}+\beta_k-2\gamma_k}{2\sqrt{\beta_k^2-3\alpha_k\gamma_k}+f'(x_{k-1})-f'(x_k)},
\end{aligned}\tag{2.3.20}$$

由于对称性，不妨设 $x_{k-1}<x_k$，且 $f'(x_{k-1})>0>f'(x_k)$. 因为

$$\beta_k^2-3\alpha_k\gamma_k=(\beta_k-\gamma_k)^2-f'(x_k)f'(x_{k-1})>0.\tag{2.3.21}$$

(2.3.20)式的右端项分母必大于零. 于是得到迭代公式

$$\begin{aligned}
x_{k+1}&=x_k-\frac{[\sqrt{(\beta_k-\gamma_k)^2-f'(x_{k-1})f'(x_k)}+\beta_k-2\gamma_k](x_k-x_{k-1})}{2\sqrt{(\beta_k-\gamma_k)^2-f'(x_{k-1})f'(x_k)}+f'(x_{k-1})-f'(x_k)}\\
&=x_k+\frac{f'(x_k)(x_k-x_{k-1})}{\sqrt{(\beta_k-\gamma_k)^2-f'(x_{k-1})f'(x_k)}-\beta_k}.
\end{aligned}\tag{2.3.22}$$

假定 $x_k\neq x_{k-1}$ 且充分靠近 x^*, $f'(x^*)=0$, $f''(x^*)>0$. 则从(2.3.16)～(2.3.17)可知

$$\alpha_k=O(|x_k-x_{k-1}|^2),\tag{2.3.23}$$

$$\beta_k=\frac{1}{2}f''(x_k)(x_k-x_{k-1})+O(|x_k-x_{k-1}|^2).\tag{2.3.24}$$

故有

$$\frac{x_{k+1}-x_k}{\hat{x}_{k+1}-x_k}=1+o(1),\tag{2.3.25}$$

其中 $\hat{x}_{k+1}$ 是由迭代公式(2.2.34)得到的. 由于迭代法(2.2.34)是 Q-超线性收敛的，所以对于迭代公式(2.3.22)也有

$$|x_{k+1}-x^*|=o(|x_k-x^*|).\tag{2.3.26}$$

为了估计(2.3.22)的精确收敛速度，我们给出如下引理：

引理 2.3.3 设 $\psi(x)\in C^4[0,\ 1]$, $\theta_H\psi(x)$ 是 $\psi(x)$ 的三次 Hermite 插值函数，满足对 $i=0,\ 1$，有

$$\theta_H\psi^{(i)}(0)=\psi^{(i)}(0),\ \theta_H\psi^{(i)}(1)=\psi^{(i)}(1).\tag{2.3.27}$$

则对充分小的 $x>0$ 有 $\xi_x\in(0,\ 1)$使得

$$\psi'(x)-\theta_H\psi'(x)=\frac{f^{(4)}(\xi_x)}{12}x(1-x)(1-2x).\tag{2.3.28}$$

证明 由 Peano 核定理

$$\psi'(x)-\theta_H\psi'(x)=\frac{1}{3!}\int_0^1 f^{(4)}(t)\frac{d}{dx}[(x-t)_+^3-\theta_H(x-t)_+^3]dt$$

$$=\frac{1}{6}\int_0^1 f^{(4)}(t)[3(x-t)_+^2-6(1-t)^3x(1-x)$$

$$-3(1-t)^2x(3x-2)]dt$$

$$=\frac{1}{6}\int_0^1 f^{(4)}(t)\bar{\psi}(x,\ t)dt. \tag{2.3.29}$$

因为 $x>0$ 且充分小，从而有：当 $0<t<x$ 时

$$\bar{\psi}(x,\ t)=3t^2[(3-2x)x^2-2(2-t)x+1]>0; \tag{2.3.30}$$

当 $x<t<1$ 时，有

$$\bar{\psi}(x,\ t)=3(1-t)^2x[2t(1-x)-x]$$

$$>3(1-t)^2x[x-2x^2]>0. \tag{2.3.31}$$

于是利用积分中值定理可得

$$\psi'(x)-\theta_H\psi'(x)=\frac{f^{(4)}(\xi)}{6}\int_0^1\bar{\psi}(x;\ t)dt$$

$$=\frac{f^{(4)}(\xi)}{6}\frac{1}{2}x(1-x)(1-2x). \tag{2.3.32}$$

于是引理成立. ▮

现在考虑

$$\psi_k(t)=f[x_k+t(x_{k-1}-x_k)], \tag{2.3.33}$$

由超线性收敛性知 $t_k=(x_{k+1}-x_k)/(x_{k-1}-x_k)\to 0$. 注意到

$$\theta_H\psi_k'(t_k)=0, \tag{2.3.34}$$

即由引理 2.3.3 可知：存在 $\xi_k\in(0,\ 1)$ 使得

$$\psi_k'(t_k)=\frac{\psi_k^{(4)}(\xi_k)}{12}t_k(1-t_k)(1-2t_k), \tag{2.3.35}$$

于是

$$f'(x_{k+1})=\frac{f^{(4)}(\hat{\xi}_k)}{12}(x_{k+1}-x_k)(x_{k-1}-x_{k+1})(x_k+x_{k-1}-2x_{k+1}), \tag{2.3.36}$$

$$x_{k+1}-x^*=\left[-\frac{f^{(4)}(x^*)}{12f''(x^*)}+O(|x_{k-1}-x^*|)\right](x_k-x^*)(x_{k-1}-x^*)^2. \tag{2.3.37}$$

由引理2.2.2即知，如果$f^{(4)}(x^*)\neq 0$，则迭代法(2.3.22)是Q-平方收敛的.

下面的算法可找到三次插值法所需的两个初始值：

算法2.3.4

步1　给定$x_1\in\mathbf{R}$，$h>0$；$\alpha>1$；

如果$f'(x_1)=0$，则停；

如果$f'(x_1)>0$，则$h:=-h$.

步2　$x_2=x_1+h$；如果$f'(x_1)f'(x_2)\leqslant 0$，则停.

$h:=2h$，$x_1=x_2$，转步2.

显然可见，上面的算法除在已求到稳定点外必产生两个点x_1、x_2满足(2.3.11). 现在我们可给出三次插值法的计算步骤：

算法2.3.5（三次插值法）

步1　给定初始$x_1\in\mathbf{R}$，$\varepsilon>0$，$k=2$

由算法2.3.4产生x_1，x_2；如果$f'(x_1)f'(x_2)=0$，则停.

步2　由迭代公式(2.3.22)计算x_{k+1}；

如果$|f'(x_{k+1})|\leqslant\varepsilon$，则停.

步3　如果$f'(x_k)f'(x_{k+1})<0$，则转步4

$x_k:=x_{k-1}$.

步4　$k:=k+1$，转步2.

从(2.3.37)可看到，如果$f^{(4)}(x^*)>0$. 则当k充分大时，步3中的条件$f'(x_k)f'(x_{k-1})<0$必成立，于是利用(2.3.37)和引理2.2.2可知点列x_k是Q-平方收敛的. 如果$f^{(4)}(x^*)<0$，则必存在k_0，当$k\geqslant k_0$时步3的条件$f'(x_k)f'(x_{k+1})<0$不成立. 故用归纳法可证在$k\geqslant k_0$次迭代时有$x_{k-1}=x_{k_0-1}$，于是从(2.3.37)式可知点列x_k仅是线性收敛且

$$\lim_{k\to\infty}|x_{k+1}-x^*|/|x_k-x^*|=\left|\frac{f^{(4)}(x^*)}{12f''(x^*)}(x_{k_0-1}-x^*)^2\right|. \tag{2.3.38}$$

§2.4 区间分割法

区间分割法是一类缩短包含极值点区间的方法. 它通过计算及比较函数在区间内某些点的值得到较小的包含极值点的区间. 这类方法最初是对单峰函数提出来的. 下面, 我们给出单峰函数的定义:

定义 2.4.1 设实值函数 $f(x)$ 在 $[a, b]$ 上有定义, 如果存在 $x^* \in [a, b]$, 使得 $f(x)$ 在 $[a, x^*]$ 上严格单调下降和在 $[x^*, b]$ 上严格单调上升, 则称 $f(x)$ 是区间 $[a, b]$ 上的**单峰函数**.

由单峰函数的定义, 引出以下引理:

引理 2.4.2 设 $f(x)$ 是 $[a, b]$ 上的单峰函数, x^* 是 $f(x)$ 在 $[a, b]$ 上唯一极小点, 则对任何 $a' \in [a, x^*]$, $b' \in [x^*, b]$, $f(x)$ 也是区间 $[a', b']$ 上的单峰函数.

引理 2.4.3 设 $f(x)$ 是 $[a, b]$ 上的单峰函数, x^* 是 $f(x)$ 在 $[a, b]$ 上的唯一极小点, 对任何 x_1、$x_2 \in [a, b]$ 满足 $x_1 < x_2$, 如果 $f(x_1) < f(x_2)$, 则 $x^* \in [a, x_2]$; 如果 $f(x_1) > f(x_2)$, 则 $x^* \in [x_1, b]$; 如果 $f(x_1) = f(x_2)$, 则 $x^* \in [x_1, x_2]$.

区间分割法就是利用引理 2.4.2 和引理 2.4.3 使包含极小点的区间通过分割不断缩小, 直至达到所需精度为止.

假设 $f(x)$ 是 $[a, b]$ 区间上的单峰函数, 任取 x_1、$x_2 \in [a, b]$, $x_1 < x_2$, 令

$$[a', b'] = \begin{cases} [a, x_2], & \text{如果 } f(x_1) < f(x_2); \\ [x_1, b], & \text{如果 } f(x_1) > f(x_2); \\ [x_1, x_2], & \text{如果 } f(x_1) = f(x_2). \end{cases}$$

则知 $f(x)$ 是 $[a', b']$ 上的单峰函数, 且 $f(x)$ 在 $[a, b]$ 上的唯一极小点必属于 $[a', b']$. 新区间的长度 $b' - a'$ 满足:

$$b' - a' \leqslant \max[x_2 - a, b - x_1]. \tag{2.4.1}$$

对任何给定的 x_1、x_2, 都存在单峰函数 $f(x)$ 使得不等式(2.4.1)的两端相等. 不难看出, 不等式(2.4.1)的右端项的下确界为

$(b-a)/2$，它是当 $x_1 \to (b+a)/2$，$x_2 \to (b+a)/2$ 时的极限值．所以我们可取 $x_1=(b+a)/2-\varepsilon$，$x_2=(b+a)/2+\varepsilon$，其中 $\varepsilon>0$ 是一充分小的正数，这样在最坏的情况下，新区间 $[a', b']$ 的长度只有 $(b-a)/2+\varepsilon$．表面看来，如此选取 x_1 和 x_2 是最好的方案．如果我们仅需要缩短区间长度一次，上述方法的确是最优的．下面我们考虑一般情形下的最优分割方案．

不失一般性，假定原始区间是 $[0, 1]$，我们称 k 级最优缩小率 r_k^* 是通过 k 次缩小区间后在最坏情形下区间长度的最短可能的极限．由上面分析可知，当 $k=1$ 时，极限值为 $\gamma_1^*=1/2$，此时两个分划点均趋于 $1/2$．为了分析 $r_k^*(k\geqslant 2)$，我们引入记号 $r_k^*(\alpha)$，它是指在 $[0, 1]$ 区间上给定一个分划点 α 后，通过 k 次缩小区间在最坏情形下区间长度的最短可能的极限．由于对称性，不妨设 $\alpha\geqslant 1/2$．设另一个分划点为 β，显然 $\beta\leqslant\alpha$．不难看出，在最坏情形下，通过两次缩短后的区间长度的最小可能（在极限意义下）为 $\max[\beta, \alpha-\beta, 1-\alpha]$，所以

$$r_2^*(\alpha)=\min_{\beta}\max[\beta, \alpha-\beta, 1-\alpha]=\max\left[\frac{\alpha}{2}, 1-\alpha\right]. \tag{2.4.2}$$

现在用归纳法证明，对 $k\geqslant 2$ 都有

$$r_k^*(\alpha)=\max[\alpha/F_{k+1}, (1-\alpha)/F_k], \tag{2.4.3}$$

其中 $F_0=0$，$F_1=1$，$F_k=F_{k-1}+F_{k-2}(k\geqslant 2)$ 是著名的斐波那契(Fibonacci)数列(见 Bell, 1940 和 Fibonacci, 1202)．显然可见，式(2.4.3)在 $k=2$ 时成立．设 $k=2, \cdots, p$ 时式(2.4.3)成立，对于 $k=p+1$，给定一个初始分划点 $\alpha\geqslant\frac{1}{2}$，设第一次缩短区间所需的另一分划点为 β，显然在最优分划时应有 $1-\alpha\leqslant\beta\leqslant\alpha$．从而有

$$\begin{aligned}
r_{p+1}^*(\alpha)&=\alpha\cdot\min_{1-\alpha\leqslant\beta<\alpha} r_p^*(\max[\beta/\alpha, 1-\beta/\alpha])\\
&=\min\Big\{\min_{\substack{1-\alpha\leqslant\beta<\alpha\\ \beta>\alpha/2}}\alpha r_p^*(\beta/\alpha), \min_{\substack{1-\alpha\leqslant\beta<\alpha\\ \beta<\alpha/2}}\alpha r_p^*(1-\beta/\alpha)\Big\}
\end{aligned}$$

$$=\min\left\{\min_{\substack{1-\alpha\le\beta<\alpha\\ \beta\ge\alpha/2}}\max\left[\frac{\beta}{F_{p+1}},\ \frac{\alpha-\beta}{F_p}\right],\right.$$

$$\left.\min_{\substack{1-\alpha\le\beta<\alpha\\ \beta\le\alpha/2}}\max\left[\frac{\alpha-\beta}{F_{p+1}},\ \frac{\beta}{F_p}\right]\right\}$$

$$=\min\left\{\max\left[\frac{\alpha}{F_{p+2}},\ \frac{1-\alpha}{F_{p+1}}\right],\ \max\left[\frac{\alpha}{F_{p+2}},\ \frac{1-\alpha}{F_p}\right]\right\}$$

$$=\max\left[\frac{\alpha}{F_{p+2}},\ \frac{1-\alpha}{F_{p+1}}\right]. \qquad (2.4.4)$$

由此可知(2.4.3)对一切 $k\geqslant 2$ 均成立，且

$$r_k^*=\min_{\alpha\geqslant 1/2} r_k^*(\alpha)=r_k^*(\alpha_k)$$

$$=r_k^*(F_{k+1}/F_{k+2})=1/F_{k+2}. \qquad (2.4.5)$$

对于需要缩小区间 k 次的方法，最优分划点 $\alpha_k=F_{k+1}/F_{k+2}$. 从(2.4.4)式的推导过程可知另一分划点 β_k 应为 $1-\alpha_k=F_k/F_{k+2}$. 我们得到的是一个对称分划，所以可假定缩小一次后的区间是 $[0,\ \alpha_k]$，区间中的另一点是 β_k，而且 $\beta_k/\alpha_k=\alpha_{k-1}$ 正好是需要缩小区间 $k-1$ 次的方法的最优分划点.

基于上面的讨论，可以给出一个求单峰函数 $f(x)$ 在 $[a,\ b]$ 区间的极小值的最优区间分割法. 对于任何给定的误差允许 $\varepsilon>0$，可找到 $[a,\ b]$ 上的一个子区间，使其包含 $f(x)$ 在 $[a,\ b]$ 上的极小点且区间长度不超过 ε. 由于该方法用到的分划区间比与斐波那契数有关，故该方法也称为斐波那契法.

算法 2.4.4

步 1　$a<b$，误差界 $\varepsilon>0$ 给出；
　　取充分小正数 $\delta<\varepsilon$；
　　找最小正数 n，使得 $F_n\geqslant(b-a)/(\varepsilon-\delta)$；
　　如果 $n\geqslant 3$，则转步 2；
　　如果 $n=1$，则停；
　　$\alpha:=(a+b)/2$，计算 $\hat{f}=f(\alpha)$，转步 5；

步 2　$\alpha:=a+(b-a)F_{n-1}/F_n$
　　$\beta:=a+b-\alpha$

计算 $f(\alpha)$, $f(\beta)$; $k:=0$

步 3　$k:=k+1$; 如果 $f(\beta)>f(\alpha)$, 则转步 4

$b:=\alpha$; $\alpha:=\beta$;

如果 $n-k=2$, 则 $\hat{f}:=f(\beta)$ 和转步 5;

$\beta:=a+(b-a)F_{n-k-2}/F_{n-k}$;

计算 $f(\beta)$, 转步 3;

步 4　$a:=\beta$; $\beta:=\alpha$;

如果 $n-k=2$, 则 $\hat{f}:=f(\alpha)$ 和转步 5;

$\alpha:=a+(b-a)F_{n-k-1}/F_{n-k}$;

计算 $f(\alpha)$, 转步 3

步 5　$\beta:=\alpha-\delta$; 计算 $f(\beta)$;

如果 $f(\beta)\leqslant\hat{f}$, 则 $b:=\alpha$;

如果 $f(\beta)\geqslant\hat{f}$, 则 $a:=\beta$; 停止.

斐波那契法的分划比是随区间 $[a, b]$ 的长度与所求精度之比的变化而变化的, 每次缩小区间的分划比 F_k/F_{k+1} 都不相同, 而且需要计算斐波那契数 F_k 与 F_{k+1}. 因此, 虽然斐波那契法是最优的分割法, 但它的应用远不如下面介绍的黄金分割法广泛.

我们考虑每次缩小区间均有固定的分划比 α. 还是设原始区间是 $[0, 1]$, 类似上面的讨论. 我们可知两个分划点应为 $1-\alpha$, α. 由于对称性, 可假定缩短后的区间是 $[0, \alpha]$. 为了使 $1-\alpha$ 仍是一个分划点, 我们要求

$$(1-\alpha)/\alpha=\alpha. \tag{2.4.6}$$

不难求出, (2.4.6)的唯一正根是 $\alpha=1/\tau$, 其中 $\tau=(\sqrt{5}+1)/2$ 在 2.2 节中已经提到.

分划比确定后, 我们可给出黄金分割法如下:

算法 2.4.5 (黄金分割法)

步 1　$a<b$, 误差界 $\varepsilon>0$ 给出;

计算 $\tau=(\sqrt{5}+1)/2$;

$\alpha:=a+(b-a)/\tau$; $\beta:=a+b-\alpha$;

计算 $f(\alpha)$, $f(\beta)$.

步 2　如果 $b-a<\tau\varepsilon$，则转步 4；

如果 $f(\beta)>f(\alpha)$，则转步 3；

$b:=\alpha$；$\alpha:=\beta$；

$\beta:=b-(b-a)/\tau$；

计算 $f(\beta)$，转步 2

步 3　$a:=\beta$；$\beta:=\alpha$；

$\alpha:=a+(b-a)/\tau$

计算 $f(\alpha)$，转步 2.

步 4　如果 $f(\beta)\leqslant f(\alpha)$，则 $b:=\alpha$；

如果 $f(\beta)\geqslant f(\alpha)$，则 $a:=\beta$；

停止.

我们显然可以看到，对任何 $k\geqslant1$，黄金分割法的 k 级缩小率(即原始区间[0, 1]通过 k 次缩小区间后的区间长度)为 $1/\tau^k$. 对于斐波那契法，我们可知(2.4.5)式在 $k=1$ 时也成立. 由于斐波那契数满足恒等式

$$F_k=[\tau^k-(-\tau)^{-k}]/\sqrt{5} \tag{2.4.7}$$

(见 Coxeter, 1954)，我们可证明下面不等式

$$\tau^{k+1}>F_{k+2}\geqslant\tau^k \tag{2.4.8}$$

对一切 $k\geqslant0$ 均成立，而且

$$\lim_{k\to\infty}F_k/F_{k+1}=1/\tau. \tag{2.4.9}$$

由此可知，黄金分割法是斐波那契法的逼近及极限情形. (2.4.8)式还表明黄金分割法的 k 级缩小率小于 $k-1$ 级最优缩小率. 换句话说，只要多计算一次函数值，黄金分割法就可以达到斐波那契法的精度. 由于黄金分割法不需要计算斐波那契数，所以它的应用远比斐波那契法广泛.

黄金分割比 $1/\tau$ 可溯源到 Euclid 的几何原本，但"黄金分割"这一名词是由 Helmes(1844)最早使用的(见 Archibald, 1918).

上面介绍的黄金分割法和斐波那契法都是只利用函数值的方法. 我们将给出一个利用函数导数值的区间分割法. 该方法每

次迭代都将区间分成相等长度的两半，所以它被称为对分法(bisection method). 对分法实质上是一个求 $f'(x)=0$ 的根的区间分割法. 设函数 $f(x)$ 在闭区间 $[a, b]$ 上可微，且 $f'(a)f'(b)<0$，则 $f'(x)=0$ 在 (a, b) 中必有根，对任何 $\alpha\in(a, b)$，如果 $f'(\alpha)=0$，则 α 即是所求的解. 否则，我们有 $f'(\alpha)f'(a)<0$ 或者 $f'(\alpha)f'(b)<0$. 令

$$[a', b']=\begin{cases}[a, \alpha], \text{ 如果 } f'(\alpha)f(a)<0;\\ [\alpha, b], \text{ 如果 } f'(\alpha)f(b)<0.\end{cases}$$

则知在 $[a', b']$ 上 $f'(x)=0$ 有根且 $[a', b']$ 区间的长度满足

$$b'-a'\leqslant\max\{\alpha-a, b-\alpha\}. \tag{2.4.10}$$

为了使新区间 $[a', b']$ 的长度在最坏的情况下达到最小，α 必须取为区间中点 $(a+b)/2$. 这时 $b'-a'=(b-a)/2$. 重复上述过程，可使区间缩小到任意给定的误差界之内. 下面给出对分法的算法.

算法 2.4.6(对分法)

步 1 $a<b$, $f'(a)f'(b)<0$ 给出, $\varepsilon>0$ 给出.

步 2 如果 $b-a<\varepsilon$, 则停.
　　$\alpha:=(a+b)/2$; 计算 $f'(\alpha)$;
　　如果 $f'(\alpha)=0$, 则停;
　　如果 $f'(\alpha)f'(a)<0$, 则转步 3;
　　$a:=\alpha$; 转步 2.

步 3 $b:=\alpha$; 转步 2.

§2.5 线 搜 索

线搜索(Line search)是对于一个多维函数在一个一维子空间上寻找一个"好"点. 在数学上，它是求解或者非精确地求解 n 维函数 $f(x)$ 在一个一维子空间上的极小，即

$$\min_{\alpha>0} f(x+\alpha d), \tag{2.5.1}$$

其中 $x\in\mathbb{R}^n$, $d\in\mathbb{R}^n$, 且 $d^T\nabla f(x)<0$. 由于大多数非线性规划方

法在每次迭代都要进行线搜索，所以线搜索方法的好坏直接影响非线性规划方法的效率.

精确线搜索(exact line search)是指精确地求解(2.5.1)，即计算 $\alpha^*>0$ 使得

$$f(x+\alpha^* d)=\min_{\alpha>0} f(x+\alpha d). \tag{2.5.2}$$

我们用 $\langle d, -\nabla f(x)\rangle$ 表示向量 d 与 $-\nabla f(x)$ 之间的角度，则有

$$\cos\langle d, -\nabla f(x)\rangle=-d^T\nabla f(x)/\|d\|_2\|\nabla f(x)\|_2. \tag{2.5.3}$$

引理 2.5.1 设 $\alpha^*>0$ 是 (2.5.2) 的解，$\|\nabla^2 f(x+\alpha d)\|_2\leqslant M$ 对一切 $\alpha>0$ 均成立，则有

$$f(x)-f(x+\alpha^* d)\geqslant\frac{1}{2M}\|\nabla f(x)\|_2^2\cos^2\langle d, -\nabla f(x)\rangle. \tag{2.5.4}$$

证明 由假设可知

$$f(x+\alpha d)\leqslant f(x)+\alpha d^T\nabla f(x)+\frac{\alpha^2}{2}M\|d\|_2^2 \tag{2.5.5}$$

对一切 $\alpha>0$ 都成立. 令 $\bar{\alpha}=-d^T\nabla f(x)/M\|d\|_2^2$，则有

$$\begin{aligned}f(x)-f(x+\alpha^* d)&\geqslant f(x)-f(x+\bar{\alpha} d)\\&=\frac{1}{2M}\|\nabla f(x)\|_2^2\cos^2\langle d, -\nabla f(x)\rangle.\end{aligned} \tag{2.5.6}$$

所以引理成立. ∎

(2.5.4)式给出在精确搜索下目标函数下降的一个下界. 它说明，只要搜索方向不是太坏，即 d 与 $-\nabla f(x)$ 的夹角不靠近 $\pi/2$，则目标函数下降在量级上至少也有 $\|\nabla f(x)\|_2^2$. 精确线搜索的另一个优点是它具有正交条件：

$$d^T\nabla f(x+\alpha^* d)=0. \tag{2.5.7}$$

这一点对一些多维无约束优化方法的有限终止起着关键作用.

精确线搜索是一个特殊形式的一维优化问题，它可用本章以上几节的方法求解. 显然可见，计算单参数函数 $\varphi(\alpha)=f(x+\alpha d)$ 的函数值和导数值实际上需要计算 n 维函数 $f(x)$ 的函数值和梯度. 所以当 n 非常大时，精确求解 (2.5.1) 的计算量是相当大

的. 由于线搜索只是多维优化方法每次迭代中的一个子问题, 精确求解它往往需要计算很多次函数值, 这有可能降低整个方法的效率. 特别是当迭代点离原问题的解很远时, 精确地求解一个一维子问题通常不是十分有效的.

非精确线搜索(inexact line search)是指不精确地求解(2.5.1), 即计算 $\alpha>0$ 使得某些判别条件满足. 非精确线搜索有不少种,但其基本思想都一样,就是使得某种类似(2.5.4)的"充分下降"条件得到满足.

常用的一种非精确线搜索, 也称 Armijo-Goldstein 线搜索(Armijo, 1966; Goldstein, 1965),要求 α 满足

$$f(x)-f(x+\alpha d)\geqslant-\alpha b_1 d^T\nabla f(x), \tag{2.5.8}$$

$$d^T\nabla f(x+\alpha d)\geqslant b_2 d^T\nabla f(x), \tag{2.5.9}$$

其中 b_1 和 b_2 是两个常数, 且满足 $0<b_1\leqslant b_2<1$ (见 Powell, 1976a). 首先我们证明除了特殊情形外这一非精确线搜索有解.

引理 2.5.2 如果函数 $\varphi(\alpha)=f(x+\alpha d)$ 连续可微, $\varphi'(0)=d^T\nabla f(x)<0$, 以及 $\varphi(\alpha)(\alpha>0)$ 下方有界, 则必有 $\alpha>0$, 使得(2.5.8)~(2.5.9)成立.

证明 由于 $\varphi(\alpha)(\alpha>0)$ 下方有界,必存在 $\hat{\alpha}>0$ 使得

$$f(x)-f(x+\hat{\alpha}d)=-\hat{\alpha}b_1 d^T\nabla f(x), \tag{2.5.10}$$

且(2.5.8)对一切 $0<\alpha\leqslant\hat{\alpha}$ 都成立. 利用中值定理和(2.5.10)可知,存在 $\bar{\alpha}\in(0,\hat{\alpha})$,使得

$$\begin{aligned}\varphi'(\bar{\alpha})&=[\varphi(0)-\varphi(\hat{\alpha})]/(-\hat{\alpha})\\&=b_1 d^T\nabla f(x).\end{aligned} \tag{2.5.11}$$

显然, $\bar{\alpha}$ 满足(2.5.8)~(2.5.9). ∎

我们可以理解条件(2.5.8)是(2.5.4)的推广,不等式(2.5.9)是(2.5.7)的近似. 所以有的非精确搜索要求(2.5.8)和

$$|d^T\nabla f(x+\alpha d)|\leqslant-b_2 d^T\nabla f(x). \tag{2.5.12}$$

这时,当 $b_2\to0$ 时就得到精确线搜索(见 Fletcher, 1987).

引理 2.5.3 设函数 $f(x)$ 连续可微, 且 $\nabla f(x)$ 满足 Lipschitz 条件

$$\|\nabla f(y)-\nabla f(x)\|_2\leqslant M\|y-x\|_2, \qquad (2.5.13)$$

如果 $f(x+\alpha d)(\alpha>0)$ 下方有界，则对任何 $\alpha>0$ 满足 (2.5.8)～(2.5.9)均有

$$f(x)-f(x+\alpha d)\geqslant\frac{b_1(1-b_2)}{M}\|\nabla f(x)\|_2^2\cos^2\langle d,\ -\nabla f(x)\rangle. \qquad (2.5.14)$$

证明 由(2.5.13)和(2.5.9)可知

$$\begin{aligned}\alpha M\|d\|_2^2&\geqslant d^T[\nabla f(x+\alpha d)-\nabla f(x)]\\&\geqslant-(1-b_2)d^T\nabla f(x),\end{aligned} \qquad (2.5.15)$$

利用(2.5.8)和(2.5.15)即可知(2.5.14)成立. ∎

求解非精确线搜索(2.5.8)～(2.5.9)的一个计算框图可见于 Fletcher(1987). 下面给出一个求满足 (2.5.8) 和 (2.5.12) 的 α 的计算方法，它的基本思想和 Fletcher 的方法一样，即利用二次插值和三次插值.

算法 2.5.4

步 1 给出 $0<b_1\leqslant b_2<1$, $\alpha:=1$;
$\alpha_1:=0$, $f_1:=f(x)$, $f_1'=d^T\nabla f(x)$;
$\alpha_2:=+\infty$, $f_2'=-1$.

步 2 计算 $f:=f(x+\alpha d)$, 如果(2.5.8)式成立则转步 4;
$\alpha_2:=\alpha$, $f_2:=f$.

步 3 利用 f_1, f_1', f_2 进行二次插值求 α, 即

$$\alpha=\alpha_1+\frac{1}{2}(\alpha_2-\alpha_1)\Big/\left[1+\frac{f_1-f_2}{(\alpha_2-\alpha_1)f_1'}\right] \qquad (2.5.16)$$

转步 2.

步 4 计算 $f'=d^T\nabla f(x+\alpha d)$, 如果(2.5.12)式成立，则停;
如果 $f'<0$, 则转步 6;
$\alpha_2:=\alpha$; $f_2:=f$, $f_2':=f'$;

步 5 利用 f_1, f_1', f_2, f_2' 进行三次插值求 α:

$$\alpha=\alpha_1-\frac{f_1'(\alpha_2-\alpha_1)}{\sqrt{(\beta-f_1')^2-f_1'f_2'}-\beta}, \qquad (2.5.17)$$

其中 $\beta=2f_1'+f_2'-3(f_2-f_1)/(\alpha_2-\alpha_1)$; 转步 2.

步 6　如果 $\alpha_2=+\infty$, 则转步 7.

$\alpha_1:=\alpha$, $f_1:=f$, $f_1':=f'$;

如果 $f_2'>0$, 则转步 5, 否则, 转步 3.

步 7　利用 f_1, f_1', f, f' 进行三次插值求 $\hat{\alpha}$:

$$\hat{\alpha}=\alpha-\frac{f'(\alpha-\alpha_1)}{\sqrt{(\hat{\beta}-f')^2-f'f_1'}+\hat{\beta}}, \tag{2.5.18}$$

其中 $\hat{\beta}=2f'+f_1'-3(f_1-f)/(\alpha_1-\alpha)$;

$\alpha_1:=\alpha$, $f_1:=f$, $f_1':=f'$;

$\alpha:=\hat{\alpha}$, 转步 2.

迭代公式(2.5.16)和(2.5.17)分别由式(2.2.29)和(2.3.22)导出. 式(2.5.18)与式(2.5.17)稍不一样, 这是因为在步 5 中我们要求得到插值函数在$[\alpha_1, \alpha_2]$上的极小点, 而在步 7 中我们是求插值函数在$(\alpha, +\infty)$上的极小点.

和 Fletcher(1987)的处理方式类似, 在算法的步 3 和步 5 中, 我们要求

$$\alpha\in[\alpha_1+\tau(\alpha_2-\alpha_1), \alpha_2-\tau(\alpha_2-\alpha_1)] \tag{2.5.19}$$

其中 $\tau>0$(通常是令 $\tau=0.1$). 如果通过(2.5.16)或者(2.5.17)计算得到的 α 不满足(2.5.19), 我们就令:

$$\alpha:=\min\{\max[\alpha, \alpha_1+\tau(\alpha_2-\alpha_1)], \alpha_2-\tau(\alpha_2-\alpha_1)\} \tag{2.5.20}$$

这样重新定义的 α 显然满足(2.5.19), 而且它是区间$[\alpha_1+\tau(\alpha_2-\alpha_1), \alpha_2-\tau(\alpha_2-\alpha_1)]$上离计算得到的 α 最近的一点. 同样在步 7 中, 我们要求

$$\hat{\alpha}\in[\alpha+(\alpha-\alpha_1), \alpha+9(\alpha-\alpha_1)]. \tag{2.5.21}$$

如果由(2.5.18)计算得到的 $\hat{\alpha}$ 不满足(2.5.21), 我们可用与(2.5.20)相同的技巧修正 $\hat{\alpha}$ 使其满足(2.5.21).

对于一致凸函数, 精确线搜索与非精确线搜索均可得到与步长有关的函数下降量的估计式. 我们设 $f(x)$是一致凸函数, 即存在常数 $\eta>0$, 使得

$$(y-z)^T[\nabla f(y)-\nabla f(z)]\geq\eta\|y-z\|_2^2 \tag{2.5.22}$$

对任何 $y、z\in\mathbf{R}^n$ 均成立. 首先对精确线搜索我们有如下结果:

定理 2.5.5 设 $\alpha^*>0$ 是 (2.5.2) 的解, 函数 $f(x)$ 满足 (2.5.22), 则必有

$$f(x)-f(x+\alpha^*d)\geqslant\frac{1}{2}\eta\|\alpha^*d\|_2^2. \qquad (2.5.23)$$

证明 由 $d^T\nabla f(x+\alpha^*d)=0$ 与(2.5.22)可得

$$\begin{aligned}f(x)-f(x+\alpha^*d)&=\int_0^{\alpha^*}-d^T\nabla f(x+td)dt\\&=\int_0^{\alpha^*}d^T[\nabla f(x+\alpha^*d)-\nabla f(x+td)]dt\\&\geqslant\int_0^{\alpha^*}\eta(\alpha^*-t)dt\|d\|_2^2\\&=\frac{1}{2}\eta\|\alpha^*d\|_2^2.\end{aligned} \qquad (2.5.24)$$

于是定理成立. ∎

对于非精确搜索, 也有类似结果:

定理 2.5.6 设 $\alpha>0$ 满足非精确线搜索条件(2.5.8), 如果函数 $f(x)$ 满足(2.5.13)和(2.5.22), 则必有

$$f(x)-f(x+\alpha d)\geqslant\frac{b_1\eta}{1+\sqrt{M/m}}\|\alpha d\|_2^2. \qquad (2.5.25)$$

证明 先假定 $d^T\nabla f(x+\alpha d)\leqslant 0$, 这时与(2.5.24)类似, 我们可证

$$f(x)-f(x+\alpha d)\geqslant\frac{1}{2}\eta\|\alpha d\|_2^2. \qquad (2.5.26)$$

现假定 $d^T\nabla f(x+\alpha d)>0$, 则必存在 $0<\alpha^*<\alpha$, 使得 $d^T\nabla f(x+\alpha^*d)=0$. 由(2.5.13)我们可证

$$f(x)-f(x+\alpha^*d)\leqslant\frac{1}{2}M\|\alpha^*d\|_2^2. \qquad (2.5.27)$$

另一方面, 利用(2.5.22)可得到

$$f(x+\alpha d)-f(x+\alpha^*d)\geqslant\frac{1}{2}m(\alpha-\alpha^*)^2\|d\|_2^2. \qquad (2.5.28)$$

由于 $f(x+\alpha d)<f(x)$, 从(2.2.27)和(2.2.28)可推出

$$\alpha\leqslant\left(1+\sqrt{\frac{M}{m}}\right)\alpha^*. \qquad (2.5.29)$$

从而有

$$
\begin{aligned}
f(x)-f(x+\alpha d) &\geqslant -\alpha b_1 d^T \nabla f(x) \\
&= b_1 \alpha d^T [\nabla f(x+\alpha^* d)-\nabla f(x)] \\
&\geqslant \eta b_1 \alpha \alpha^* \|d\|_2^2 \\
&\geqslant \frac{b_1 \eta}{1+\sqrt{M/m}} \|\alpha d\|_2^2.
\end{aligned}
\tag{2.5.30}
$$

由于 $b_1<1$, $M/m>1$, 从(2.5.30)及(2.5.26)即知定理成立。

第 3 章

梯度法和共轭梯度法

本章讨论极小化问题

$$\min_{x\in\mathbf{R}^n} f(x) \tag{3.0.1}$$

的梯度法和共轭梯度法．梯度法是多维无约束优化方法中最简单的也是最基本的．它具有简单、可靠等优点，但缺点是收敛较慢．梯度法中的步长选取对算法的影响十分大，对这个问题还需进行更深入的研究．共轭梯度法是利用负梯度方向和已有的搜索方向产生新的搜索方向．其基本思想是使得这些方向在$f(x)$是二次凸函数时相互共轭．共轭梯度法是从解线性方程组的共轭梯度法发展而来的，它具有二次终止性．但对于一般非线性函数$f(x)$，共轭梯度法的收敛速度仅为线性，所以实用的共轭梯度法都要加上重新开始技巧．共轭梯度法是解大规模优化问题的一类主要方法．

§3.1 梯　度　法

梯度法(Cradient Methods)是以负梯度方向作为搜索方向的方法，是所有需要计算导数的无约束优化算法中最简单的方法．

设函数$f(x)$在x点附近连续可微，且 $g=\nabla f(x)\neq 0$．不难看到，负梯度方向$-g/\|g\|$是一个下降方向，即对充分小的$\alpha>0$有

$$f(x-\alpha g/\|g\|)<f(x). \tag{3.1.1}$$

对任何非零向量 $d\in\mathbf{R}^n$ 均有

$$\lim_{\alpha\to0_+}\frac{f(x)-f(x+\alpha d/\|d\|_2)}{f(x)-f(x-\alpha g/\|g\|_2)}=-\frac{d^Tg}{\|d\|_2\|g\|_2}$$
$$=\cos\langle d-g\rangle\leqslant1. \tag{3.1.2}$$

其中 $\langle d, -g\rangle$ 表示向量 d 与 $-g$ 之间的夹角. 从 (3.1.2) 式可知, 负梯度方向 $-g/\|g\|_2$ 是函数 $f(x)$ 在 x 点唯一的最速下降方向 (steepest descent direction), 所以梯度法也被称为最速下降法. 经典的最速下降法每次迭代都用精确线搜索求步长因子, 这方法可溯源到 Cauchy(1847). 它的计算步骤如下:

算法 3.1.1 (最速下降法)

步 1 给出 $x_1\in\mathbf{R}^n$, $0\leqslant\varepsilon\ll1$;

$k:=1$.

步 2 如果 $\|\nabla f(x_k)\|_2\leqslant\varepsilon$, 则停;

$d_k=-\nabla f(x_k)$.

步 3 利用精确线搜索求 $\alpha_k>0$. 即

$$f(x_k+\alpha_kd_k)=\min_{\alpha>0}f(x_k+\alpha d_k). \tag{3.1.3}$$

步 4 $x_{k+1}=x_k+\alpha_kd_k$, $k:=k+1$, 转步 2.

很显然, 算法 3.1.1 是一个下降算法, 即满足

$$f(x_{k+1})<f(x_k). \tag{3.1.4}$$

有如下收敛性定理:

定理 3.1.2 设函数 $f(x)$ 二次连续可微, 且 $\|\nabla^2f(x)\|\leqslant M$. 对任何给定的初始值 x_1 和 $\varepsilon>0$, 算法 3.1.1 必有限终止或者

$$\lim_{k\to\infty}f(x_k)=-\infty. \tag{3.1.5}$$

证明 我们考虑无穷迭代下去的算法 3.1.1. 由引理 2.5.1 有

$$f(x_k)-f(x_{k+1})\geqslant\frac{1}{2M}\|\nabla f(x_k)\|_2^2. \tag{3.1.6}$$

于是

$$f(x_1)-f(x_k)=\sum_{i=1}^{k-1}[f(x_i)-f(x_{i+1})]$$
$$\geqslant\frac{1}{2M}\sum_{i=1}^{k-1}\|\nabla f(x_i)\|_2^2. \tag{3.1.7}$$

从(3.1.7)可知(3.1.5)与

$$\lim_{k\to\infty}\|\nabla f(x_k)\|_2=0 \tag{3.1.8}$$

至少有一个成立. 从而定理成立. ▌

下面我们分析算法 3.1.1 的收敛速度. 假定 $\varepsilon=0$, 算法 3.1.1 产生的点列 $x_k(k=1, 2, \cdots)$满足以下假设:

假设 3.1.3 1) $x_k\to x^*$;

2) $\nabla f(x^*)=0$, $\nabla^2 f(x^*)$正定.

分析无约束优化算法的收敛速度常常要用到目标函数 $f(x)$ 的下列性质:

引理 3.1.4 设 $f(x)$在 x^* 点附近二次连续可微, $\nabla f(x^*)=0$, $\nabla^2 f(x^*)$正定. 则存在正数 m, M, δ, 使得

$$0<m\leqslant\sigma_n(\nabla^2 f(x))\leqslant\sigma_1(\nabla^2 f(x))\leqslant M \tag{3.1.9}$$

对一切$\|x-x^*\|\leqslant\delta$ 均成立. 其中 $\sigma_1(B)$, $\sigma_n(B)$分别表示 $n\times n$ 矩阵 B 的最大特征值和最小特征值. 关系式

$$\frac{1}{2}m\|x-x^*\|_2^2\leqslant f(x)-f(x^*)\leqslant\frac{1}{2}M\|x-x^*\|_2^2, \tag{3.1.10}$$

$$m\|x-x^*\|_2\leqslant\|\nabla f(x)\|\leqslant M\|x-x^*\|_2. \tag{3.1.11}$$

对一切 $\|x-x^*\|\leqslant\delta$ 均成立. 而且对任何充分小的 $m'\in(0, m)$, 存在 $\delta'>0$ 使得

$$\|\nabla f(x)\|_2^2\geqslant 2(m-m')[f(x)-f(x^*)] \tag{3.1.12}$$

对一切 $\|x-x^*\|\leqslant\delta'$ 均成立.

证明 (3.1.9)～(3.1.11)是十分显然的. 如果引理不对, 则必存在 $m'\in(0, m)$, $y_k\to x^*$, 使得

$$\|\nabla f(y_k)\|_2^2<2(m-m')[f(y_k)-f(x^*)]. \tag{3.1.13}$$

不失一般性, 假定$(y_k-x^*)/\|y_k-x^*\|_2\to\bar{d}$, 我们有

$$\|\nabla f(y_k)\|_2^2=\|y_k-x^*\|^2(\|\nabla^2 f(x^*)\bar{d}\|_2^2+o(1)), \tag{3.1.14}$$

$$f(y_k)-f(x^*)=\frac{1}{2}\|y_k-x^*\|^2(\bar{d}^T\nabla^2 f(x^*)\bar{d}+o(1)). \tag{3.1.15}$$

由(3.1.14)～(3.1.15)以及(3.1.9)可得

$$\lim_{k\to\infty}\frac{\|\nabla f(y_k)\|_2^2}{f(y_k)-f(x^*)}=\frac{2\|\nabla^2 f(x^*)\bar{d}\|_2^2}{\bar{d}^T\nabla^2 f(x^*)\bar{d}}\geqslant m. \tag{3.1.16}$$

(3.1.16)与(3.1.13)矛盾. 此矛盾说明引理成立. ▌

下面定理给出算法 3.1.1 的收敛速度.

定理3.1.5 设由算法 3.1.1 产生的点列满足假设 3.1.3, 设 $f(x)$在 x^* 处附近二次连续可微, 且(3.1.9)式成立, 则必有

$$\frac{f(x_{k+1})-f(x^*)}{f(x_k)-f(x^*)}=\beta_k<1, \tag{3.1.17}$$

且

$$\lim_{k\to+\infty}\sup\beta_k\leqslant\frac{M-m}{M}<1. \tag{3.1.18}$$

证明 由(3.1.6)式有

$$\beta_k\leqslant 1-\frac{\|\nabla f(x_k)\|_2^2}{2M[f(x_k)-f(x^*)]}<1. \tag{3.1.19}$$

类似于证明(3.1.16), 有

$$\begin{aligned}\lim_{k\to\infty}\sup\beta_k&\leqslant 1-\lim_{k\to\infty}\text{nif}\ \frac{\|\nabla f(x_k)\|_2^2}{2M[f(x_k)-f(x^*)]}\\&\leqslant 1-\frac{m}{M}<1.\end{aligned} \tag{3.1.20}$$

所以定理成立. ▌

如果$f(x)$是二次严格凸函数, 即

$$f(x)=g^Tx+\frac{1}{2}x^THx, \tag{3.1.21}$$

其中 $g\in\mathbf{R}^n$, $H\in\mathbf{R}^{n\times n}$ 对称正定, 显然唯一的极小值点是 $x^*=-H^{-1}g$. 记 $\sigma_1(H)$、$\sigma_n(H)$分别为H的最大特征值和最小特征值, 则由算法 3.1.1 产生的点列 $x_k(k=1, 2, \cdots)$满足

$$\frac{f(x_{k+1})-f(x^*)}{f(x_k)-f(x^*)}\leqslant\left[\frac{\sigma_1(H)-\sigma_n(H)}{\sigma_1(H)+\sigma_n(H)}\right]^2. \tag{3.1.22}$$

上式的证明可参阅 Akaike(1959). 而且 Forsythe(1968)还证明了: 如果 $x_2\neq x^*$, 则对一切 k 均有 $x_k\neq x^*$, 且存在常数 $c>0$, 使得

$$\frac{f(x_{k+1})-f(x^*)}{f(x_k)-f(x^*)}\geqslant c>0 \tag{3.1.23}$$

对任何 k 均成立. 更有趣的是, 当 k 充分大时 $g_k=\nabla f(x_k)$ 将在两个方向上来回摆动, 即

$$\lim_{k\to\infty} g_{2k}/\|g_{2k}\|_2=\bar{d}, \tag{3.1.24}$$

$$\lim_{k\to\infty} g_{2k+1}/\|g_{2k+1}\|_2=\hat{d} \tag{3.1.25}$$

都存在. 由(2.1.23)和(3.1.10)可知, 即使目标函数是二次函数, 算法 3.1.1 也仅是线性收敛. 而且从下例可以看出, 当矩阵 H 的条件数相当大时, 算法 3.1.1 的收敛是相当慢的.

考虑

$$f(x)=\frac{1}{2}x^T\begin{bmatrix}M & 0\\ 0 & m\end{bmatrix}x \tag{3.1.26}$$

其中 $x\in\mathbf{R}^2$, $M>1>m>0$. 设初始值为 $x_1=(1,\ \sqrt{M/m})$, 则由算法 3.1.1 产生的点列是

$$x_k=\left(\frac{M-m}{M+m}\right)^{k-1}((-1)^{k-1},\ \sqrt{M/m}), \tag{3.1.27}$$

当 $M\gg m$ 时, $(M-m)/(M+m)$ 十分靠近 1, 从(3.1.27)可知 x_k 收敛于解(0, 0)的速度是线性的, 且非常慢.

最近, Barzilai 和 Borwein(1988)提出一个两点步长梯度法, 其基本思想是利用迭代当前点以及前一点的信息来确定步长因子. 他们把迭代公式 $x_{k+1}=x_k-\alpha_k g_k$ 看成是

$$x_{k+1}=x_k-D_k g_k, \tag{3.1.28}$$

其中 $g_k=g(x_k)=\nabla f(x_k)$, $D_k=\alpha_k I$. 为了使矩阵 D_k 具有"拟牛顿"(quasi-Newton)性质(拟牛顿方法是一类重要的优化方法, 我们将在下一章专门讨论). 计算 α_k 使得

$$\min\|s_{k-1}-D_k y_{k-1}\|_2 \tag{3.1.29}$$

或者

$$\min\|D_k^{-1}s_{k-1}-y_{k-1}\|_2, \tag{3.1.30}$$

其中 $s_{k-1}=x_k-x_{k-1}$, $y_{k-1}=g_k-g_{k-1}$. 由(3.1.29)和(3.1.30)我们可分别求得

$$\alpha_k=s_{k-1}^T y_{k-1}/\|y_{k-1}\|_2^2 \tag{3.1.31}$$

和

$$\alpha_k=\|s_{k-1}\|_2^2/s_{k-1}^T y_{k-1}. \tag{3.1.32}$$

利用步长因子(3.1.31)或(3.1.32)，得到下面算法：

算法3.1.6 (Barzilar and Borwein, 1988)

步1 给出 $x_1\in\mathbf{R}^n$, $0\leqslant\varepsilon\ll 1$;

$k:=1$

步2 如果$\|\nabla f(x_k)\|_2\leqslant\varepsilon$，则停；

$$d_k=-\nabla f(x_k).$$

步3 如果 $k=1$ 则利用线搜索求 α_1，否则，

由 (3.1.31)或(3.1.32)给出 α_k：

步4 $x_{k+1}=x_k+\alpha_k d_k$, $k:=k+1$，转步 2.

我们所考虑的目标函数是两个变量的二次函数时算法3.1.6的收敛速度，由于对变量作正交变换不影响算法，故可假定目标函数是

$$f(x)=\frac{1}{2}x^T\begin{bmatrix}1 & 0\\ 0 & \lambda\end{bmatrix}x, \tag{3.1.33}$$

其中 $\lambda\geqslant 1$. 对于初始值 $x_1=(\lambda,\ \lambda^{-1})$，$\alpha_1$ 由精确线性搜索求得，Barzilar 和 Borwein (1988)证明了步长由 (3.1.31) 确定的算法3.1.6 是 R-超线性收敛的，而且 R-阶为$\sqrt{2}$. 这个结果几乎对所有初始值都是成立的.

定理3.1.7 设算法3.1.6中的步长 α_k 由(3.1.32)给出. 如果初始值可表示为

$$x_1=t_1(\lambda,\ \pm 1)^T, \tag{3.1.34}$$

且 α_1 由精确搜索给出，即

$$(g(x_2))^T g(x_1)=0, \tag{3.1.35}$$

则 x_k 线性收敛于(0, 0)，且

$$x_k=t_1\left(\frac{\lambda-1}{\lambda+1}\right)^{k-1}(\lambda,\ \pm(-1)^{k-1})^T; \tag{3.1.36}$$

如果(3.1.34)或者(3.1.35)不成立，则 x_k R-超收敛于解(0, 0)，且 R-阶为 $\sqrt{2}$.

证明 如果(3.1.34)和(3.1.35)成立，则有

$$\alpha_1=\frac{\|g(x_1)\|_2^2}{g(x_1)^T\begin{bmatrix}1 & 0\\ 0 & \lambda\end{bmatrix}g(x_1)}$$

$$=\frac{2}{1+\lambda}. \tag{3.1.37}$$

从而有

$$x_2=x_1-\alpha_1 g(x_1)$$

$$=t_1\left(\frac{\lambda-1}{\lambda+1}\right)(\lambda,\ \mp 1)^T. \tag{3.1.38}$$

利用归纳法即可证(3.1.36)对一切 k 都成立.

现在假定(3.1.34)和(3.1.35)不同时成立,设 $x_1=(u_1,\ v_1)^T$, $x_2=(u_2,\ v_2)^T$,如果 x_1 或者 x_2 落在 u-轴上或 v-轴上,则算法最多经过两次迭代就可得到精确解. 所以我们可以假定 $u_1,\ v_1,\ u_2,\ v_2$ 都不为零,于是我们可将 g_1、g_2 表示成如下形式:

$$g_1=t_1(\lambda^{m_1},\ (-1)^{I_1})^T, \tag{3.1.39}$$

$$g_2=t_2(\lambda^{m_2},\ (-1)^{I_2})^T, \tag{3.1.40}$$

其中 I_1, $I_2\in\{0,\ 1\}$. 假定

$$g_i=t_i(\lambda^{m_i},\ (-1)^{I_i})^T \tag{3.1.41}$$

对一切 $i=1,\ 2,\ \cdots,\ k$ 均成立,我们从(3.1.32)可得到

$$\alpha_k=\|s_{k-1}\|_2^2/s_{k-1}^T\begin{bmatrix}1 & 0\\ 0 & \lambda\end{bmatrix}s_{k-1}$$

$$=(\lambda^{2m_{k-1}}+1)/(\lambda^{2m_{k-1}}+\lambda). \tag{3.1.42}$$

于是

$$g_{k+1}=g_k-\alpha_k\begin{bmatrix}1 & 0\\ 0 & \lambda\end{bmatrix}g_k$$

$$=t_k\frac{(\lambda-1)\lambda^{2m_{k-1}}}{\lambda^{2m_{k-1}}+\lambda}\begin{bmatrix}\lambda^{m_k-2m_{k-1}}\\ (-1)^{I_k+1}\end{bmatrix}. \tag{3.1.43}$$

对于 $k\geqslant 2$, 我们定义

$$m_{k+1}=m_k-2m_{k-1}, \tag{3.1.44}$$

$$t_{k+1}=t_k(\lambda-1)\lambda^{2m_{k-1}}/(\lambda^{2m_{k-1}}+\lambda), \tag{3.1.45}$$

$$I_{k+1}=I_k+1. \tag{3.1.46}$$

则由(3.1.39)、(3.1.40)、(3.1.43)可知关系式(3.1.41)对一切 $i\geqslant 1$ 均成立. 从(3.1.44)可知(见 Hardy 和 Wright, 1971)存在常数 θ、ϕ 使得

$$m_k=(\sqrt{2})^k\theta\cos(\phi+k\tan^{-1}(\sqrt{7})) \tag{3.1.47}$$

对一切 k 都成立. 由于(3.1.34)和(3.1.35)不都成立, 所以 $\theta\neq 0$. 不失一般性, 假定 $\theta>0$, 对任何 k, 都有

$$\begin{aligned}&\min_{i=0,1,2,3}\cos(\phi+(k+i)\tan^{-1}\sqrt{7})\\ &\leqslant\cos\left(\frac{3}{2}\tan^{-1}\sqrt{7}\right)\leqslant -0.24.\end{aligned} \tag{3.1.48}$$

从(3.1.45)可知

$$|t_{k+1}|\leqslant(\lambda-1)|t_k| \tag{3.1.49}$$

对所有 $k\geqslant 2$ 都成立, 而且如果 $m_{k-1}<0$, 则有

$$|t_{k+1}|\leqslant(\lambda-1)\lambda^{2m_{k-1}-1}|t_k|. \tag{3.1.50}$$

由(3.1.45)、(3.1.47)～(3.1.50)可证

$$\begin{aligned}|t_k|&=\left[\prod_{i=1}^{k-2}(\lambda-1)\frac{\lambda^{2m_i}}{\lambda^{2m_i}+\lambda}\right]|t_2|\\ &\leqslant|t_2|(\lambda-1)^{k-2}\prod_{i=1}^{[(k-2)/4]}\lambda^{-0.48\theta(\sqrt{2})^{k-1-4i}}\\ &\leqslant|t_2|(\lambda-1)^{k-2}\lambda^{-0.48\theta(\sqrt{2})^{k-5}}\end{aligned} \tag{3.1.51}$$

对所有的 $k\geqslant 6$ 均成立. 于是当 $k\geqslant 6$ 且 $m_k<0$ 时,

$$\begin{aligned}\|g_k\|_2&\leqslant\sqrt{2}\,|t_k|\\ &\leqslant\sqrt{2}\,|t_2|(\lambda-1)^{k-2}\lambda^{-0.48\theta(\sqrt{2})^{k-5}}.\end{aligned} \tag{3.1.52}$$

从(3.1.45)和(3.1.41)可得

$$\begin{aligned}\|g_{k+1}\|_2^2/\|g_k\|_2^2&=\left(\frac{(\lambda-1)\lambda^{2m_{k-1}}}{\lambda^{2m_{k-1}}+\lambda}\right)^2\frac{\lambda^{2m_{k+1}}+1}{\lambda^{2m_k}+1}\\ &=(\lambda-1)^2\frac{\lambda^{2m_k}+\lambda^{4m_{k-1}}}{(\lambda^{2m_{k-1}}+\lambda)^2(\lambda^{2m_k}+1)}\\ &\leqslant(\lambda-1)^2\end{aligned} \tag{3.1.53}$$

对一切 $k\geqslant 2$ 均成立. 所以对任何 $k\geqslant 9$, 由(3.1.48)知, 必存在 $0\leqslant j\leqslant 3$ 使得 $m_{k-j}<0$. 利用(3.1.53)和(3.1.52)可证

$$\|g_k\|_2 \leqslant (\lambda-1)^j \|g_{k-j}\|_2$$
$$\leqslant (\lambda-1)^j \sqrt{2}\,|t_2|\,(\lambda-1)^{k-j-2}\lambda^{-0.4889(\sqrt{2})^{k-j-3}}$$
$$\leqslant \sqrt{2}\,|t_2|\,(\lambda-1)^{k-2}\lambda^{-0.4889(\sqrt{2})^{k-3}}. \qquad (3.1.54)$$

由(3.1.54)即知 $\|g_k\|_2$ 是 R-超线性收敛于 0 的，且具有收敛阶 $\sqrt{2}$. 由于 $\|x_k\|_2 \leqslant \|g_k\|_2$，即知 x_k R-超线性收敛于解(0, 0)，且 R-收敛阶为 $\sqrt{2}$. ▮

定理 3.1.7 在步长 α_k 由(3.1.31)给出的情形下仍然成立，而且证明十分类似. 从不等式(3.1.54)可看出算法 3.1.6 的一个有趣性质，当问题越是处于条件不利时，即 λ 越大时，$\|g_k\|$ 收敛于 0 就越快.

算法 3.1.6 并不是一个下降方法，所以当目标函数不是二次函数时，要加以适当修改才可应用. 从定理 3.1.7 的证明中可看到，当靠近解时，每 4 次迭代必有一次超线性收敛步. 于是我们可以修正算法，要求每 4 次迭代目标函数值一定下降.

算法 3.1.8

步 1　给出 $x_1 \in \mathbf{R}^n$, $0 \leqslant \varepsilon \ll 1$, $0 < b_1 \leqslant b_2 < 1$;
　　　$k := 1$, index $:= 1$, $f_{\min} = f(x_1)$.

步 2　如果 $\|\nabla f(x_k)\|_2 \leqslant \varepsilon$，则停;
　　　$d_k = -\nabla f(x_k)$.

步 3　如果 $k = 1$ 或者 $k -$ index $= 4$，则转步 4，由(3.1.31)或(3.1.32)计算 α_k，如果 $\alpha_k \leqslant 0$ 或无定义，则转步 4；转步 5.

步 4　$x_k := x_{\text{index}}$; index $:= k$;
　　　利用非精确搜索(2.5.8)与(2.5.9)求 $\alpha_k > 0$;

步 5　$x_{k+1} = x_k + \alpha_k d_k$, $k := k+1$;
　　　如果 $f(x_k) \geqslant f_{\min}$，则转步 2;
　　　index $:= k$; $f_{\min} = f(x_k)$; 转步 2.

在算法 3.1.8 中，如果只有有限次通过步 4，则该算法对充分大的 k 和算法 3.1.6 完全一样，且保证每 4 次迭代必缩小目标函数值. 如果有无穷次通过步 4 求 α_k，则有

$$f(x_k)-f(x_{k_i+1})\geqslant\frac{b_1(1-b_2)}{M}\|g_{k_i}\|_2^2 \tag{3.1.55}$$

对一切 $i=1, 2, \cdots$, 成立. 其中 k_i 是第 i 次通过步 4 求 α_k 时 k 的值. 由算法可知

$$f(x_{k_i})=\min_{1\leqslant j\leqslant k_i}f(x_j). \tag{3.1.56}$$

显然 $k_{i+1}\geqslant k_i+1$, 由(3.1.56)即知

$$f(x_{k_{i+1}})\leqslant f(x_{k_i+1}). \tag{3.1.57}$$

所以算法产生的点列 x_k 满足

$$\lim_{k\to+\infty}\inf f(x_k)=-\infty \tag{3.1.58}$$

或者

$$\lim_{k\to\infty}\inf\|\nabla f(x_k)\|_2=0. \tag{3.1.59}$$

§ 3.2　共轭梯度法

共轭梯度法(Conjugate Gradient Methods)是利用目标函数梯度逐步产生共轭方向作为线搜索方向的方法.

考虑目标函数是凸的二次函数的情形, 即

$$f(x)=g^Tx+\frac{1}{2}x^THx, \tag{3.2.1}$$

其中 $g\in\mathbf{R}^n$, $H\in\mathbf{R}^{n\times n}$ 对称正定. 共轭方向的定义如下:

定义 3.2.1　设 $H\in\mathbf{R}^{n\times n}$ 对称正定, 对 $\mathbf{R}^n$ 中的任何一组非零向量 $d_1, d_2, \cdots, d_m$, 如果

$$d_i^THd_j=0\quad(i\neq j), \tag{3.2.2}$$

则称 $d_1, d_2, \cdots, d_m$ 相互 H-共轭.

显然可见, 如果 $d_1, \cdots, d_m$ 相互 H-共轭, 则 $d_1, \cdots, d_m$ 必线性无关. 假设我们通过某种方法得到了几个相互 H-共轭的向量 $d_1, d_2, \cdots, d_n$, 由于 $d_1, \cdots, d_n$ 线性无关, 我们可把 x 表示为

$$x=\sum_{i=1}^{n}\alpha_id_i. \tag{3.2.3}$$

将(3.2.3)代入(3.2.1)，利用共轭性质(3.2.2)，我们可把求 $f(x)$ 极小转化为

$$\min_{\alpha_i \in \mathbf{R}^n} \sum_{i=1}^{n} \left(\alpha_i d_i^T g + \frac{1}{2} \alpha_i^2 d_i^T H d_i \right). \tag{3.2.4}$$

显而易见，(3.2.2)的解为

$$\alpha_i = -d_i^T g / d_i^T H d_i \quad (i=1,\ 2,\ \cdots,\ n). \tag{3.2.5}$$

由于确定这些参数 $\alpha_i(i=1,\ 2,\ \cdots,\ n)$ 都是相互独立的，我们并不要求一开始就有所有的 n 个 H-共轭方向. 例如：假定一开始只有一个方向 d_1，我们可利用(3.2.5)求出 α_1，然后利用某种方式得到一个与 d_1 H-共轭的方向 d_2，再利用(3.2.5)计算 α_2，依此下去，直到把 α_n 求出. 共轭梯度法正是按这种方式产生搜索方向 d_i.

共轭梯度法一般取第一个搜索方向为负梯度方向，即

$$d_1 = -g_1 = -Hx_1 + g, \tag{3.2.6}$$

然后通过精确线搜索求出步长 α_1：

$$\alpha_1 = \|g_1\|_2^2 / g_1^T H g_1 \tag{3.2.7}$$

和

$$x_2 = x_1 + \alpha_1 d_1.$$

由于 α_1 是由精确线搜索给出的，故必有

$$g_2^T d_1 = 0. \tag{3.2.8}$$

所以可取

$$d_2 = -g_2 + \beta_1 d_1 \tag{3.2.9}$$

使得 d_2 与 d_1 H-共轭. 不难求出

$$\beta_1 = \frac{g_2^T H d_1}{d_1^T H d_1}, \tag{3.2.10}$$

对任何 $k \geqslant 1$，共轭梯度法取

$$d_{k+1} = -g_{k+1} + \beta_k d_k + \sum_{i=1}^{k-1} \beta_k^{(i)} d_i, \tag{3.2.11}$$

使得 d_{k+1} 与 $d_1,\ \cdots,\ d_k$ H-共轭. 利用归纳法可证

$$g_{k+1}^T d_i = 0 \quad (i=1,\ 2,\ \cdots,\ k), \tag{3.2.12}$$

所以，只要 $g_{k+1} \neq 0$，则可由(3.2.11)求出一个与 $d_1,\ \cdots,\ d_k$ H-共轭的下降方向，事实上

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|_2^2. \tag{3.2.13}$$

由(3.2.11)可知 g_i 是 $d_1, \cdots, d_i$ 的线性组合，所以从(3.2.12)可得到

$$g_{k+1}^T g_i = 0 \quad (i=1, 2, \cdots, k), \tag{3.2.14}$$

于是

$$\begin{aligned} g_{k+1}^T H d_i &= g_{k+1}^T [g_{i+1} - g_i] \\ &= 0 \quad (i=1, 2, \cdots, k-1). \end{aligned} \tag{3.2.15}$$

在等式(3.2.11)的两端左乘 $d_i^T H (i=1, \cdots, k)$，就得到

$$\beta_k^{(i)} = 0, \tag{3.2.16}$$

$$\beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k}. \tag{3.2.17}$$

现在我们可将共轭梯度法给出如下算法：

算法 3.2.2 （二次函数的共轭梯度法）

步 1　给出初值 $x_1 \in \mathbf{R}^n$，如果 $\|g_1\| = 0$，则停；

$d_1 = -g_1$, $k=1$；

步 2　计算精确搜索步长

$$\alpha_k = -g_k^T d_k / d_k^T H d_k \tag{3.2.18}$$

$x_{k+1} = x_k + \alpha_k d_k$,

如果 $\|g_{k+1}\|_2 = 0$，则停；

由(3.2.17)计算 β_k；

$$d_{k+1} = -g_{k+1} + \beta_k d_k \tag{3.2.19}$$

$k := k+1$，转步 2.

由于 $d_1, \cdots, d_k$ 相互 H-共轭，从(3.2.12)可知必存在 $m \leqslant n$，使得

$$g_{m+1} = 0. \tag{3.2.20}$$

于是我们已得到了如下结果：

定理 3.2.3　算法 3.2.2 经过不超过 n 次迭代就会终止，即存在 $m \leqslant n$，使得

$$g_{m+1} = 0, \tag{3.2.21}$$

而且对一切 $1 \leqslant k \leqslant m$，都有

$$g_k^T d_k = -\|g_k\|_2^2, \tag{3.2.22}$$

$$d_k^T H d_j = 0, \tag{3.2.23}$$

$$g_k^T d_j = 0, \tag{3.2.24}$$

$$g_k^T g_j = 0, \tag{3.2.25}$$

其中 $j=1, 2, \cdots, k-1$.

证明 由 d_k 的定义即知(3.2.23)成立. 公式(3.2.21)、(3.2.22)、(3.2.24)和(3.2.25)分别由(3.2.20)、(3.2.13)、(3.2.12)和(3.2.14)给出. ∎

利用定理 3.2.3 可知, 当函数 $f(x)$ 是(3.2.1)时, 下列公式

$$\beta_k = \|g_{k+1}\|_2^2 / \|g_k\|_2^2, \tag{3.2.26}$$

$$\beta_k = (g_{k+1} - g_k)^T g_{k+1} / \|g_k\|_2^2, \tag{3.2.27}$$

$$\beta_k = -\|g_{k+1}\|_2^2 / d_k^T g_k, \tag{3.2.28}$$

都和 (3.2.17) 是等价的. 公式 (3.2.26) 和 (3.2.27) 分别称为 Fletcher-Reeves 公式和 Polak-Ribiere-Polyak 公式(参阅 Fletcher 和 Reeves, 1964, Polak 和 Ribiére, 1969; 以及 Polyak, 1969). 公式(3.2.28)称为共轭下降公式(见 Fletcher, 1987). 显然, 当目标函数是二次的情况下, 在算法 3.2.2 中的 β_k 用(3.2.26)～(3.2.28) 中任何一个式子求都产生一相同的点列. 但是对于一般的非线性函数, (3.2.26)～(3.2.28)并不相互等价, 从而导出不同方法. 下面给出的算法是基于(3.2.26)的方法.

算法 3.2.4(Fletcher-Reeves 方法)

步 1 给出 $x_1 \in \mathbf{R}^n$, $0 \leqslant \varepsilon < 1$.

$d_1 = -g_1 = \nabla f(x_1)$, $k := 1$;

步 2 如果 $\|g_k\|_2 \leqslant \varepsilon$, 则停;

利用某种线搜索方法求 α_k;

$x_{k+1} = x_k + \alpha_k d_k$.

步 3 利用(3.2.26)计算 β_k;

$d_{k+1} = -g_{k+1} + \beta_k d_k$;

$k := k+1$; 转步 2.

如果把步 3 中利用(3.2.26)计算 β_k 换成利用 (3.2.27) 计算

β_k, 则得到 Polak-Ribiére-Polyak 方法, 简称为 PRP 方法. 下面我们考虑算法 3.2.4 的收敛性.

设目标函数 $f(x)$ 二次连续可微, 且由算法 3.2.4 产生的点列 x_k 有界. $\alpha_k>0$ 由精确线搜索或者非精确线搜索求得, 由引理 2.5.1 和引理 2.5.3, 我们可知, 必存在常数 $\delta>0$, 使得

$$f(x_k)-f(x_k+\alpha_k d_k)\geqslant-\delta(d_k^T g_k)^2/\|d_k\|^2. \qquad (3.2.29)$$

定理 3.2.5 设 $f(x)$ 二次连续可微, $\{x_k\}$ 由算法 3.2.4 产生且有界. $\alpha_k>0$ 由某种线性搜索求得满足 (3.2.29), 且存在常数 $b_2<\frac{1}{2}$, 使得

$$|d_k^T\nabla f(x_k+\alpha_k d_k)|\leqslant-b_2 d_k^T\nabla f(x_k) \qquad (3.2.30)$$

对一切 k 都成立, 则必有

$$\liminf_{k\to\infty}\|g_k\|_2=0. \qquad (3.2.31)$$

证明 首先用归纳法证明

$$|d_k^T g_k+\|g_k\|_2^2|\leqslant u\|g_k\|_2^2. \qquad (3.2.32)$$

其中 $u=b_2/(1-b_2)<1$. (3.2.32) 显然对 $k=1$ 成立, 假定 (3.2.32) 对 $k=1, 2, \cdots, i$ 时都成立, 则有

$$\begin{aligned}|d_{i+1}^T g_{i+1}+\|g_{i+1}\|_2^2| &= \frac{\|g_{i+1}\|_2^2}{\|g_i\|_2^2}|g_{i+1}^T d_i| \\ &\leqslant b_2\|g_{i+1}\|_2^2\frac{|g_i^T d_i|}{\|g_i\|_2^2} \\ &\leqslant b_2\|g_{i+1}\|_2^2(1+u) \\ &= u\|g_{i+1}\|_2^2.\end{aligned} \qquad (3.2.33)$$

所以 (3.2.32) 对一切 k 都成立. 由 (3.2.26) 可得

$$\begin{aligned}\|d_{k+1}\|_2^2 &= \|g_{k+1}\|_2^2+\frac{\|g_{k+1}\|_2^4}{\|g_k\|_2^4}\|d_k\|_2^2-2\frac{\|g_{k+1}\|^2}{\|g_k\|_2^2}g_{k+1}^T d_k \\ &\leqslant \|g_{k+1}\|_2^2+\frac{\|g_{k+1}\|_2^4}{\|g_k\|_2^4}\|d_k\|_2^2+2u\|g_{k+1}\|_2^2.\end{aligned} \qquad (3.2.34)$$

令 $t_k=\|d_k\|_2^2/\|g_k\|_2^4$, 则有

$$t_{k+1}\leqslant t_k+\frac{1+2u}{\|g_{k+1}\|_2^2}. \qquad (3.2.35)$$

如果定理不真，则存在 $\delta>0$，使得

$$\|g_k\|_2^2\geqslant\delta' \tag{3.2.36}$$

对一切 k 均成立. (3.2.35)和(3.2.36)表明

$$t_k\leqslant\frac{1+2u}{\delta'}k+t_1. \tag{3.2.37}$$

另一方面，由于点列 $\{x_k\}$ 有界，从而有

$$\begin{aligned}\infty&>\sum_{k=1}^{\infty}[f(x_k)-f(x_{k+1})]\\&\geqslant\sum_{k=1}^{\infty}-\delta(d_k^Tg_k)^2/\|d_k\|_2^2\\&\geqslant\sum_{k=1}^{\infty}\delta(1-u)^2\|g_k\|_2^4/\|d_k\|_2^2\\&\geqslant\sum_{k=1}^{\infty}\frac{\delta(1-u)^2}{t_k}.\end{aligned} \tag{3.2.38}$$

由于点列 $\{x_k\}$ 有界，则 $\|g_k\|_2$ 必上方有界，从(3.2.38)可知级数

$$\sum_{k=1}^{\infty}1/t_k<+\infty \tag{3.2.39}$$

必收敛，这与(3.2.37)式矛盾. 此矛盾说明了定理是正确的. ▮

Powell(1984a) 在精确线搜索的假定下 (即 $b_2=0$) 证明了定理3.2.5. Al-Baali(1985) 将 Powell 结果推广到非精确线搜索 ($b_2<1/2$). 至今，对于 $b_2\in\left[\frac{1}{2},\ 1\right)$，非精确线搜索的 Fletcher-Reeves 方法是否收敛仍是一个迷.

对于另一个重要的共轭梯度法 PRP 方法(即 β_k 由(3.2.27)给出)，Powell(1984a)给出了一个有趣的例子：在精确线搜索下的 PRP 方法产生的点列 x_k 在 6 点附近循环. 这 6 点中任何一点都不是目标函数的稳定点. 这个例子的发现的确有点令人意外，因为在实际计算中，PRP 方法比 Fletcher-Reeves 方法好得多(见 Powell, 1977). 由于 Powell 的例子，如果我们只限定目标函数二次连续可微，那么无论是精确线搜索还是非精确线搜索的 PRP 方法均可能产生一串点列 $x_k(k=1,\ 2,\ \cdots)$使得

$$\liminf_{k\to\infty}\|\nabla f(x_k)\|_2>0. \tag{3.2.40}$$

现在考虑共轭下降法(即 β_k 由(3.2.28)给出). 如果每次迭代 α_k 均由精确线搜索给出,则

$$-d_k^T g_k=\|g_k\|_2^2. \tag{3.2.41}$$

于是共轭下降法和 Fletcher-Reeves 方法一样, 必有(3.2.31)成立. 对于非精确搜索的情形,有以下定理:

定理 3.2.6 设 $f(x)$ 二次连续可微, $\{x_k\}$ 由共轭下降法产生有界, $\alpha_k>0$ 由某种线搜索求得满足(3.2.29), 且存在常数 $b_2<\frac{1}{2}$, 使得

$$0\geqslant d^T\nabla f(x_k+\alpha_k d_k)\geqslant b_2 d_k^T\nabla f(x_k) \tag{3.2.42}$$

对一切 k 都成立,则必有

$$\liminf_{k\to\infty}\|\nabla f(x_k)\|_2=0. \tag{3.2.43}$$

证明 与定理 3.2.5 的证明类似,首先,我们有

$$0>d_k^T g_k+\|g_k\|_2^2=\beta_{k-1}g_k^T d_{k-1}\geqslant -b_2\|g_k\|_2^2, \tag{3.2.44}$$

于是

$$\begin{aligned}\|d_{k+1}\|_2^2&=\|g_{k+1}\|_2^2+\frac{\|g_{k+1}\|_2^4}{(d_k^T g_k)^2}\|d_k\|_2^2-2\frac{\|g_{k+1}\|_2^2}{g_k^T d_k}g_{k+1}^T d_k\\&\leqslant\|g_{k+1}\|_2^2+\frac{\|g_{k+1}\|_2^4}{\|g_k\|_2^4}\|d_k\|_2^2.\end{aligned} \tag{3.2.45}$$

利用(3.2.44)和(3.2.45), 同定理 3.2.5 的证明几乎一样, 可以证明(3.2.43)必成立. ∎

尽管 Powell 的例子说明了 PRP 方法对一般非线性函数可能不收敛,但对于一致凸函数,有如下结果:

定理 3.2.7 设 $f(x)$ 二次连续可微, 且一致凸, 即存在常数 $\eta>0$, 使得对任何 y、$z\in\mathbf{R}^n$, 均有

$$(y-z)^T[\nabla f(y)-\nabla f(z)]\geqslant\eta\|y-z\|_2^2. \tag{3.2.46}$$

假定 $\alpha_k>0$, 由精确线搜索得到或者满足(2.5.8)与(2.5.9), 则由 PRP 方法产生的点列 $\{x_k\}$ 必有

$$\lim_{k\to\infty}\|\nabla f(x_k)\|_2=0. \tag{3.2.47}$$

证明 假定定理不真,则存在 $\delta>0$, 使得

$$\|g_k\|_2 \geqslant \delta \tag{3.2.48}$$

对一切 k 都成立. 由引理 2.5.5 和引理 2.5.6 可知存在常数 $\eta' > 0$, 使得

$$f(x_k) - f(x_{k+1}) \geqslant \eta' \|x_k - x_{k+1}\|_2^2. \tag{3.2.49}$$

由于 $f(x)$ 一致凸, 故 $f(x)$ 必下方有界, 于是

$$\sum_{k=1}^{\infty} \|x_k - x_{k+1}\|_2^2 \leqslant \frac{1}{\eta'} \sum_{k=1}^{\infty} [f(x_k) - f(x_{k+1})] < +\infty, \tag{3.2.50}$$

从而有

$$|\beta_k| = |g_{k+1}^T [g_{k+1} - g_k] / \|g_k\|^2| \to 0. \tag{3.2.51}$$

由于函数 $f(x)$ 的一致凸性, 点列 $\{x_k\}$ 必有界, 于是 $\|g_k\|$ 有界, 利用

$$\|d_{k+1}\| \leqslant \|g_{k+1}\| + |\beta_k| \, \|d_k\| \tag{3.2.52}$$

可证 $\|d_k\|$ 必有界. 对于精确搜索, 有

$$d_k^T g_k = -\|g_k\|_2^2. \tag{3.2.53}$$

对于阿米久型非精确搜索, 我们有(当 k 充分大时)

$$|d_k^T g_k| = |-\|g_k\|_2^2 + \beta_k g_k^T d_{k-1}| \geqslant \frac{1}{2} \|g_k\|_2^2. \tag{3.2.54}$$

于是, 当 k 充分大时

$$\cos^2(d_k, -g_k) \geqslant \frac{1}{2} \frac{\|g_k\|_2^2}{\|d_k\|_2^2}. \tag{3.2.55}$$

因为 $\|g_k\| \geqslant \delta$, 且 $\|d_k\|$ 上方有界, 必存在 $\bar{\delta} > 0$, 使得

$$\cos^2(d_k, -g_k) \geqslant \bar{\delta}. \tag{3.2.56}$$

由于 $f(x_k)$ 下方有界, 由引理 2.5.1 和引理 2.5.2 可知

$$\sum_{k=1}^{\infty} \|g_k\|_2^2 \cos^2(d_k, -g_k) < +\infty. \tag{3.2.57}$$

另一方面, 从(3.2.56)和(3.2.48)可知(3.2.57)左端的级数应该发散. 这一矛盾说明定理正确. ▌

另一点值得提到的是: 当目标函数是(3.2.1)时, 极小化问题等价于求解线性方程组

$$Hx = g. \tag{3.2.58}$$

不难发现, 算法 3.2.2 和解线性方程组的共轭梯度法是完全等价

的．事实上，第一个求极小值的共轭梯度法的 Fletcher-Reeves 方法正是从 Hestenes 和 Stiefel(1952) 提出的解线性方程组的共轭梯度法直接发展而来的．

共轭梯度法的二次终止性是基于(3.2.6)．如果第一个搜索方向不是最速下降方向，则对严格凸的二次函数共轭梯度法一般不会有限终止(Powell, 1976b)．而且从下一节的分析可知，共轭梯度法的收敛速度仅是线性的．

Powell(1977)提出重开始(Restart)技巧，即共轭梯度法每迭代 n 步后把当前点的负梯度取为下一次迭代的搜索方向．由于每 n 步中都有一次是用最速下降方向做搜索方向，所以无论是精确线搜索还是非精确线搜索都可保证重开始共轭梯度法收敛．不同形式的共轭梯度法采用重开始技术，可使收敛速度从原来的线性提高到 n 步超线性收敛，即

$$\lim_{k\to\infty}\frac{\|x_{k+n}-x^*\|_2}{\|x_k-x^*\|_2}=0. \tag{3.2.59}$$

于是，我们看到重开始共轭梯度法无论是在总体收敛性还是在局部收敛速度方面都比原来的共轭梯度法具有更好的性质．但令人不解的是，Fletcher(1987)给出的数值结果表明重开始的 Fletcher-Reeves 方法和重开始的 PRP 方法并没有比原始的 Fletcher-Reeves 方法和 PRP 方法好多少．也就是说，重开始技术对共轭梯度法的改进没有人们想象的那么大．

关于共轭方向法的一些推广和详细分析可见 Dennis 和 Turner(1985)．

§3.3 共轭梯度法的线性收敛性

共轭梯度法具有二次终止性，但对于一个一般的非线性函数 $f(x)$，只有经过若干次迭代后，迭代点都在解的附近，我们才可把 $f(x)$ 近似地看成二次函数．即使考虑最特殊的情形，即 $f(x)$ 在解的附近是一个凸的二次函数，我们也不能应用已有的二次终止结

果，这是因为一般说来，对所有的 $k>1$，都有

$$d_k \neq -g_k. \tag{3.3.1}$$

所以，为了研究共轭梯度法的局部收敛性，我们假定目标函数是二次函数

$$f(x)=\frac{1}{2}x^T H x, \tag{3.3.2}$$

且假定初始线搜索方向 d_1 是任意一个下降方向，即只要求

$$d_1^T g_1<0. \tag{3.3.3}$$

我们假定每次迭代都是用精确搜索确定步长 α_k，于是有

$$x_{k+1}=x_k+\alpha_k d_k, \tag{3.3.4}$$

$$d_{k+1}=-g_{k+1}+\beta_k d_k, \tag{3.3.5}$$

$$g_{k+1}=g_k+\alpha_k H d_k, \tag{3.3.6}$$

$$d_{k+1}^T g_{k+1}=-\|g_{k+1}\|_2^2, \tag{3.3.7}$$

对一切 $k\geqslant 1$ 均成立，α_k 和 β_k 分别由(3.2.16)和(3.2.15)确定.

从(3.3.5)式可知 $d_{k+1}+g_{k+1}$ 可由 d_k 表示，利用精确线搜索性质和共轭方向的性质可得

$$d_{k+1}^T H(d_{k+1}+g_{k+1})=0, \tag{3.3.8}$$

$$g_{k+1}^T(d_{k+1}+g_{k+1})=0, \tag{3.3.9}$$

对一切 $k\geqslant 1$ 均成立.

引理3.3.1 如果

$$d_k+g_k=0, \tag{3.3.10}$$

则必有 $k\leqslant 2$ 或者 $g_k=d_k=0$.

证明 假定 $k>2$，由(3.3.10)和共轭性可知

$$d_{k-1}^T H g_k=-d_{k-1}^T H d_k=0. \tag{3.3.11}$$

由(3.3.6)和(3.3.11)可推出

$$d_{k-1}^T H g_{k-1}=g_{k-1}^T d_{k-1}\frac{\|Hd_{k-1}\|_2^2}{d_{k-1}^T H d_{k-1}}, \tag{3.3.12}$$

利用(3.3.12)、(3.3.8)和(3.3.9)，我们有

$$d_{k-1}^T H g_{k-1}=\|g_{k-1}\|_2^2\frac{\|Hd_{k-1}\|_2^2}{d_{k-1}^T H g_{k-1}}. \tag{3.3.13}$$

上式说明 g_{k-1} 与 Hd_{k-1} 在同一个方向，从而必有 $g_k=0$. ∎

由精确线搜索及共轭性可得:

$$d_{k-1}^T g_{k+1} = d_{k-1}^T g_k + \alpha_k d_{k-1}^T H d_k = 0. \qquad (3.3.14)$$

于是

$$g_{k+1}^T g_k = g_{k+1}^T(-d_k + \beta_{k-1} d_{k-1}) = 0 \qquad (3.3.15)$$

对一切 $k \geqslant 2$ 都成立. 现在我们给出一个类似定理 3.2.2 的结果.

引理 3.3.2 假定共轭梯度法N次迭代后终止, 即

$$g_{N+1} = 0, \qquad (3.3.16)$$

而且对 $k = N,\ N-1,\ \cdots,\ N-l+1$, 都有

$$d_k + g_k \neq 0. \qquad (3.3.17)$$

则必有

$$d_i^T H d_j = 0, \qquad (3.3.18)$$

$$g_i^T g_j = 0, \qquad (3.3.19)$$

对一切 $N-l \leqslant i < j \leqslant N$ 都成立.

证明 对于 $l=1$ 引理显然成立. 我们假定引理对 $l = l_0$ 成立. 不难看出, 要证明引理对 $l = l_0 + 1$ 成立, 只需证明(3.3.18)和(3.3.19)对 $i = N - l_0 - 1$ 和 $i < j \leqslant N$ 成立即可.

$$\begin{aligned} d_{N-l_0-1}^T H d_j &= \beta_{N-l_0-1}^{-1} (d_{N-l_0} + g_{N-l_0})^T H d_j \\ &= \beta_{N-l_0-1}^{-1} g_{N-l_0}^T H d_j \\ &= \begin{cases} -\beta_{N-l_0-1}^{-1} g_{N-l_0}^T g_N, & \text{若 } j = N \\ \alpha_j^{-1} \beta_{N-l_0-1}^{-1} g_{N-l_0}^T [g_{j+1} - g_j], & \text{若 } j < N \end{cases} \\ &= 0. \end{aligned} \qquad (3.3.20)$$

$$\begin{aligned} g_{N-l_0-1}^T g_j &= (g_{N-l_0} - \alpha_{N-l_0-1} H d_{N-l_0-1})^T g_j \\ &= -\alpha_{N-l_0-1} d_{N-l_0-1}^T H g_j \\ &= -\alpha_{N-l_0-1} d_{N-l_0-1}^T H [d_j - \beta_{j-1} d_{j-1}] \\ &= 0. \end{aligned} \qquad (3.3.21)$$

于是由归纳法即知引理成立.

利用上面的引理, 我们可得到下面一般情形下共轭梯度法的二次终止特点: 即如果有限终止, 则必终止在最初的 $n+1$ 次迭代. 如以下定理所述.

定理 3.3.3 假定目标函数是二次函数, 精确线搜索的共轭

梯度法在第N次迭代后终止，即(3.3.16)成立，则必有 $N\leqslant n+1$。

证明 假定 $N>n+1$. 令 $l=N-2$，由引理3.1.1知(3.3.17)成立. 从而有

$$g_i^T g_j=0 \tag{3.3.22}$$

对所有的 $2\leqslant i<j<N$. 由算法假定 $g_i(i=2,\cdots,N)$ 均为非零向量(否则，算法将不需要 N 次迭代就终止了). 于是(3.3.22)说明 $\mathbf{R}^n$ 中存在 $N-1>n$ 个相互正交的非零向量，而这是不可能的. 这一矛盾说明定理成立. ∎

由于

$$\begin{aligned}\frac{f(x_{k+1})-f(x^*)}{f(x_k)-f(x^*)}&=\frac{g_{k+1}^T H^{-1}g_{k+1}}{g_k^T H^{-1}g_k}\\&=1-\frac{(g_k^T d_k)^2}{d_k^T H d_k g_k^T H^{-1}g_k}\\&\leqslant 1-\frac{\sigma_n(H)}{\sigma_1(H)}<1.\end{aligned} \tag{3.3.23}$$

其中 $\sigma_1(H)$ 和 $\sigma_n(H)$ 分别是H的最大特征值和最小特征值. 我们知道共轭梯度法的收敛速度至少是线性的. 下面可证明，只要共轭梯度法不有限终止，它的收敛速度正好是线性的.

假定算法不有限终止，故必有：

$$d_k+g_k\neq 0 \tag{3.3.24}$$

对一切 k 均成立. 首先有以下引理：

引理 3.3.4

$$Hd_k=-\alpha_k^{-1}[d_{k+1}-(1+\beta_k)d_k+\beta_{k-1}d_{k-1}] \tag{3.3.25}$$

对所有的 $k>1$ 均成立.

证明 (3.3.6)和(3.3.5)可改写成

$$Hd_k=\alpha_k^{-1}[g_{k+1}-g_k], \tag{3.3.26}$$

$$g_{k+1}=-d_{k+1}+\beta_k d_k. \tag{3.3.27}$$

因为 $k>1$，将(3.3.27)中的 k 换成 $k-1$ 后得到

$$g_k=-d_k+\beta_{k-1}d_{k-1}. \tag{3.3.28}$$

用(3.3.26)～(3.3.28)即可推导出(3.3.25). ∎

引理 3.3.5 对任何整数 $l>1$，如果 d_k, d_{k+1}, $\cdots$, d_{k+l} 相互

H-共轭，则必有

$$d_j^T g_i=0,\ k\leqslant j<i\leqslant k+l+1. \tag{3.3.29}$$

证明 利用精确线搜索性质和引理中的假设，即知

$$\begin{aligned} d_j^T g_i &= d_j^T\left[g_{j+1}+\sum_{t=j+1}^{i-1}(g_{t+1}-g_t)\right] \\ &= d_j^T\left[g_{j+1}+\sum_{t=j+1}^{i-1}\alpha_t H d_t\right]=0. \end{aligned} \tag{3.3.30}$$

从而引理为真. ▌

引理 3.3.6 对任给的正整数 l，如果 $d_k,\ d_{k+1},\ \cdots,\ d_{k+l}$ 相互 H-共轭，则 $d_{k+1},\ d_{k+2},\ \cdots,\ d_{k+l+1}$ 也相互 H-共轭.

证明 显然，只需证明 d_{k+l+1} 与 $d_{k+i}(i=1,\ 2,\ \cdots,\ l)$ 共轭. 根据 d_k 的定义知 d_{k+l+1} 与 d_{k+l} H-共轭. 对于 $1\leqslant i\leqslant l-1$，我们利用(3.3.5)、(3.3.25)和(3.3.29)，可得到

$$\begin{aligned} d_{k+l+1}^T H d_j &= (-g_{k+l+1}+\beta_{k+l}d_{k+l})^T H d_j \\ &= -g_{k+l+1}^T H d_j \\ &= \alpha_j^{-1} g_{k+l+1}^T(d_{j+1}-(1+\beta_j)d_j+\beta_{j-1}d_{j-1}) \\ &= 0. \end{aligned} \tag{3.3.31}$$

于是引理成立. ▌

设 L 是使 $d_1,\ d_2,\ \cdots,\ d_L$ 相互 H-共轭的最大正数，由引理 3.3.6 可知

$$d_{L+1}^T H d_1\neq 0, \tag{3.3.32}$$

而且有以下引理:

引理 3.3.7 设 $d_1,\ d_2,\ \cdots,\ d_L$ 相互 H-共轭，且(3.3.32)成立，则对任何 $k\geqslant 1$，均有 $d_k,\ d_{k+1},\ \cdots,\ d_{k+L-1}$ 相互 H-共轭，且

$$d_{k+L}^T H d_k\neq 0. \tag{3.3.33}$$

证明 利用归纳法即知只需证明引理对 $k=2$ 成立就足够了. 从引理 3.3.6 知 $d_2,\ \cdots,\ d_{L+1}$ 相互 H-共轭. 由(3.3.25)和(3.3.29)可得到:

$$\begin{aligned} d_{L+2}^T H d_2 &= (-g_{L+2}+\beta_{L+1}d_{L+1})^T H d_2 = -g_{L+2}^T H d_2 \\ &= \alpha_2^{-1} g_{L+2}^T(d_3-(1+\beta_2)d_2+\beta_1 d_1) \end{aligned}$$

$$= \frac{\beta_1}{\alpha_2} g_{L+2}^T d_1 = \alpha_2^{-1}\beta_1(g_{L+1} + \alpha_{L+1} H d_{L+1})^T d_1$$
$$= \alpha_2^{-1}\beta_1\alpha_{L+1} d_{L+1}^T H d_1 \neq 0. \tag{3.3.34}$$

从而引理对 $k=2$ 成立. ▌

下面的引理指出比值 $\|d_k\|_2/\|g_k\|_2$ 对所有 k 上方下方均一致有界,而且指出搜索步长 α_k 也是上方下方一致有界.

引理 3.3.8 对所有的 $k>1$, 均有

$$1 \leqslant \|d_k\|_2/\|g_k\|_2 \leqslant \sqrt{K_2(H)}, \tag{3.3.35}$$

$$\frac{1}{K_2(H)\sigma_1(H)} \leqslant \alpha_k \leqslant \frac{1}{\sigma_n(H)}, \tag{3.3.36}$$

其中 $K_2(H) = \sigma_1(H)/\sigma_n(H)$, $\sigma_1(H)$、$\sigma_n(H)$ 分别为 H 的最大特征值和最小特征值.

证明 (3.3.35)的第一个不等式可由(3.3.5) 直接推出. 由(3.3.6)以及 $-g_k^T d_k = \|g_k\|_2^2$ 可得

$$\begin{aligned} 0 \leqslant \|g_{k+1}\|_2^2 &= \|g_k + \alpha_k H d_k\|_2^2 \\ &= \|g_k\|_2^2 + 2\alpha_k g_k^T H d_k + \alpha_k^2 \|H d_k\|_2^2 \\ &= \|g_k\|_2^2 \left(\frac{\|g_k\|_2^2 \|H d_k\|_2^2}{(d_k^T H d_k)^2} - 1 \right). \end{aligned} \tag{3.3.37}$$

于是

$$\begin{aligned} \|d_k\|_2^2/\|g_k\|^2 &\leqslant \|d_k\|_2^2 \|H d_k\|_2^2/(d_k^T H d_k)^2 \\ &\leqslant K_2(H). \end{aligned} \tag{3.3.38}$$

从而(3.3.35)得证. 由(3.3.9)和(3.3.35)、(3.3.36)得

$$\alpha_k = -\frac{g_k^T d_k}{d_k^T H d_k} = \frac{\|g_k\|_2^2}{d_k^T H d_k} \leqslant \frac{\|d_k\|_2^2}{d_k^T H d_k} \leqslant \frac{1}{\sigma_n(H)} \tag{3.3.39}$$

以及

$$\begin{aligned} \alpha_k = \frac{\|g_k\|_2^2}{d_k^T H d_k} &\geqslant \frac{\|d_k\|_2^2}{K_2(H) d_k^T H d_k} \\ &\geqslant \frac{1}{K_2(H)\sigma_1(H)}. \end{aligned} \tag{3.3.40}$$

从(3.3.39)、(3.3.40)即知不等式(3.3.36)成立. ▌

有了这些引理, 可以证明共轭梯度法的收敛速度正好是线性的.

定理3.3.9　设 $d_1, \cdots, d_L$ 相互 H-共轭，且(3.3.32)成立，则共轭梯度法的收敛速度仅是线性的，且对任何 k 都有：

$$\|g_{k+L+1}\|_2 \geqslant \frac{|\beta_1 g_{L+2}^T d_1|}{\|g_2\|_2^2[K_2(H)]^{2L+0.5}}\|g_k\|. \tag{3.3.41}$$

证明　由(3.3.6)和引理3.3.5可知

$$\begin{aligned} g_{L+2}^T d_1 &= [g_{L+1} + \alpha_{L+1} H d_{L+1}]^T d_1 \\ &= \alpha_{L+1} d_{L+1}^T H d_1 \neq 0. \end{aligned} \tag{3.3.42}$$

于是只需证明(3.3.41)对一切 k 都成立，则显然算法的收敛速度仅是线性的。

与(3.3.34)类似，我们可证

$$d_{L+k+1}^T H d_{k+1} = \beta_k \alpha_{k+1}^{-1} \alpha_{L+k} d_{L+k}^T H d_k \tag{3.3.43}$$

对一切 k 均成立。同样，与(3.3.42)类似

$$g_{L+k+1}^T d_k = \alpha_{L+k} d_{L+k}^T H d_k. \tag{3.3.44}$$

所以，从(3.3.43)和(3.3.44)可得递推关系

$$g_{L+k+2}^T d_{k+1} = \beta_k \alpha_{k+1}^{-1} \alpha_{L+k+1} g_{L+k+1}^T d_k. \tag{3.3.45}$$

显然 $L \geqslant 1$，由(3.2.15)、(3.2.16)和引理3.3.5可知对任何 $k>1$ 都有

$$\begin{aligned} \beta_k = \frac{g_{k+1}^T H d_k}{d_k^T H d_k} &= \frac{g_{k+1}^T [d_{k+1} - (1+\beta_k) d_k + \beta_{k-1} d_{k-1}]}{\alpha_k d_k^T H d_k} \\ &= \|g_{k+1}\|_2^2 / \|g_k\|_2^2. \end{aligned} \tag{3.3.46}$$

由(3.3.45)、(3.3.46)和(3.3.36)可得：

$$\begin{aligned} |g_{k+L+1}^T d_k| &= |g_{L+2}^T d_1| \prod_{i=1}^{k-1} |\beta_i| \alpha_{i+1}^{-1} \alpha_{L+i+1} \\ &= |\beta_1 g_{L+2}^T d_1| \frac{\|g_k\|_2^2}{\|g_2\|_2^2} \prod_{i=1}^{L} \alpha_{i+1}^{-1} \alpha_{k+i} \\ &\geqslant |\beta_1 g_{L+2}^T d_1| \frac{\|g_k\|_2^2}{\|g_2\|_2^2} [K_2(H)]^{-2L}. \end{aligned} \tag{3.3.47}$$

另一方面有

$$\begin{aligned} |g_{k+L+1}^T d_k| &\leqslant \|g_{k+L+1}\|_2 \|d_k\|_2 \\ &\leqslant \|g_{k+L+1}\|_2 \|g_k\|_2 \sqrt{K_2(H)}. \end{aligned} \tag{3.3.43}$$

从(3.3.47)和(3.3.46)可看出(3.3.41)必成立。▮

§3.4 共轭梯度法的进一步改进

考虑目标函数为一般非线性函数，共轭梯度法产生下一个搜索方向如下：

$$d_{k+1}=-g_{k+1}+\beta_k d_k. \qquad (3.4.1)$$

因为$f(x)$不是二次函数，我们不能用 $d_{k+1}Hd_k=0$ 来确定 β_k. 利用中值定理可知，存在 $t\in(0,1)$，使得

$$\alpha_k^{-1}d_{k+1}^T[g_{k+1}-g_k]=d_{k+1}^T\nabla^2 f(x_k+t\alpha_k d_k)d_k. \qquad (3.4.2)$$

所以可用条件

$$d_{k+1}^T[g_{k+1}-g_k]=0 \qquad (3.4.3)$$

来代替二次函数时的 $d_{k+1}^T H d_k=0$，由(3.4.1)和(3.4.3)可知

$$\beta_k=\frac{g_{k+1}^T y_k}{d_k^T y_k}. \qquad (3.4.4)$$

其中 $y_k=g_{k+1}-g_k$. 显然，当$f(x)$是二次函数时，(3.4.4)和(3.2.15)是等价的. Perry(1978)发现，当 β_k 由(3.4.4)式给出时，(3.4.1)可改写成如下等价形式：

$$\begin{aligned}d_{k+1}&=-\left(I-\frac{d_k y_k^T}{d_k^T y_k}\right)g_{k+1}\\&=-P_k g_{k+1}.\end{aligned} \qquad (3.4.5)$$

显然 $P_k^T y_k=0$，所以 P_k 是一个从 $\mathbf{R}^n$ 到 y_k 的零空间的仿射变换. 考虑更一般形式的(3.4.5)，即

$$d_{k+1}=-\widetilde{P}_k g_{k+1}, \qquad (3.4.6)$$

$\widetilde{P}_k$是某$n\times n$矩阵. 由于我们希望 d_{k+1} 是下降方向，故要求 $\widetilde{P}_k$ 满足

$$g_{k+1}^T\widetilde{P}_k g_{k+1}>0. \qquad (3.4.6')$$

显然可见，任何正定矩阵 $\widetilde{P}_k$ 都满足(3.4.6′). 但不足的是：正定矩阵 P 不可能满足 $P^T y_k=0$　不过，我们发现，共轭性条件(3.4.3)并不要求 $\widetilde{P}_k y_k=0$，而只需要

$$d_{k+1}^T y_k=-g_{k+1}^T\widetilde{P}_k^T y_k=0. \qquad (3.4.7)$$

在精确线搜索的假定下，只需

$$\tilde{P}_k^T y_k = \alpha_k d_k = s_k \tag{3.4.8}$$

就够了．(3.4.8)正好是“拟牛顿”方程，因为我们有

$$[\nabla^2 f(x_k)]^{-1} y_k \approx x_{k+1} - x_k = \alpha_k d_k = s_k, \tag{3.4.9}$$

如何构造 $\tilde{P}_k$ 使其满足“拟牛顿”方程(3.4.8)的方法很多，在下一章中我们将详细讨论．Shanno (1978)利用 BFGS 公式得到

$$\tilde{P}_k = I - \frac{s_k y_k^T + y_k s_k^T}{s_k^T y_k} + \left(1 + \frac{\|y_k\|_2^2}{s_k^T y_k}\right)\frac{s_k s_k^T}{s_k^T y_k}. \tag{3.4.10}$$

如果 $s_k^T y_k > 0$，则 $\tilde{P}_k$ 对称正定且满足(3.4.8)．由 (3.4.6)可得到

$$d_{k+1} = -g_{k+1} - y_k \frac{s_k^T g_{k+1}}{s_k^T y_k} + s_k \left[\left(1 + \frac{\|y_k\|_2^2}{s_k^T y_k}\right)\frac{s_k^T g_{k+1}}{s_k^T y_k} - \frac{y_k^T g_{k+1}}{s_k^T y_k}\right]. \tag{3.4.11}$$

于是我们可给出下面的算法：

算法 3.4.1

步 1　给出 $x_1 \in \mathbf{R}^n$, $\varepsilon \geqslant 0$;
　　$d_1 = -g_1$; $k := 1$;

步 2　如果 $\|g_k\| \leqslant \varepsilon$，则停；
　　利用某种线性搜索方法求 $\alpha_k > 0$;
　　$x_{k+1} = x_k + \alpha_k d_k$;

步 3　利用(3.4.11)计算 d_{k+1};
　　$k := k+1$; 转步 2.

由于 (3.4.10) 给出的 $\tilde{P}_k$ 与 $\tilde{P}_{k-1}$ 无任何关系，所以算法 3.4.1 也被称为无记忆拟牛顿方法(见 Buckley, 1982)．算法 3.4.1 比共轭梯度法(即 d_{k+1} 由 (3.4.5) 定义) 无论是在内存还是在每次迭代的计算量都没有增加多少，但它的计算表现比共轭梯度法好得多．

另一种修正的共轭梯度法是预条件(Preconditioned)共轭梯度法，它的基本思想是对一个给定的正定矩阵 A，考虑变量代换

$$\tilde{x} = A^{-\frac{1}{2}} x. \tag{3.4.12}$$

目标函数变为

$$\tilde{f}(\tilde{x})=f(x)=f(A^{\frac{1}{2}}\tilde{x}). \qquad (3.4.13)$$

令 $\tilde{x}_k=A^{-\frac{1}{2}}x_k$, $\tilde{g}_k=\nabla\tilde{f}(\tilde{x}_k)=A^{\frac{1}{2}}\nabla f(x_k)=A^{\frac{1}{2}}g_k$, 则有

$$\tilde{d}_k=A^{-\frac{1}{2}}d_k, \quad \tilde{y}_k=A^{\frac{1}{2}}y_k. \qquad (3.4.14)$$

从而公式(3.4.5)变成

$$\tilde{d}_{k+1}=-\left(A-\frac{\tilde{d}_k\tilde{y}_k^TA}{\tilde{d}_k^T\tilde{y}_k}\right)\tilde{g}_{k+1}. \qquad (3.4.15)$$

同(3.4.10)类似, 我们还可以得到公式

$$\tilde{d}_{k+1}=-\hat{P}_k\tilde{g}_{k+1}. \qquad (3.4.16)$$

其中

$$\hat{P}_k=A-\frac{\tilde{s}_k\tilde{y}_k^TA+A\tilde{y}_k\tilde{s}_k^T}{\tilde{s}_k^T\tilde{y}_k}+\left(1+\frac{\tilde{y}_k^TA\tilde{y}_k}{\tilde{s}_k^Ty_k}\right)\frac{\tilde{s}_k\tilde{s}_k}{\tilde{s}_k^T\tilde{y}_k}. \qquad (3.4.17)$$

由(3.4.16)和(3.4.17), 如果取 $A=I$, 则得到无记忆共轭梯度法; 如果 $A=\hat{P}_{k-1}$, 则得到的是 BFGS 方法. 这种把共轭梯度法看成是一种特殊的拟牛顿法, 无疑为我们对共轭梯度法的理解提供了一个有力的工具.

另一种共轭梯度法是利用负梯度以及前几步的搜索方向来确定下一步的线搜索方向, 即 d_{k+1} 具有下列形式:

$$d_{k+1}=-Ag_{k+1}+\beta_kd_k+\sum_{i=1}^{I}\beta_k^{(i)}d_{k-i}. \qquad (3.4.18)$$

$I\geqslant1$ 是某一正整数. 我们知道, 在目标函数 $f(x)$ 是二次凸函数时, 这些 $\beta_k^{(i)}$ 均为零. 对于一般非线性函数, 考虑非零 $\beta_k^{(i)}$ 有可能提高算法的效率. 例如, Beale(1972)给出的重开始共轭梯度法就用到

$$d_{k+1}=-Ag_{k+1}+\beta_kd_k+\gamma_kd_i, \qquad (3.4.19)$$

其中 i 是最后一次重开始的下标. 当 $k>i$ 时,

$$\beta_k=\frac{g_{k+1}^TAy_k}{d_k^Ty_k}, \quad \gamma_k=\frac{g_{k+1}^TAy_i}{d_i^Ty_i}. \qquad (3.4.20)$$

对于最基本的共轭梯度法(即 d_{k+1} 由 (3.4.1) 给出), 如果用

共轭性条件(3.4.3)，则 β_k 由(3.4.4)唯一确定．但是，我们考虑目标函数为二次时 β_k 的等价形式，则可得到许多不同的表达方式．如何选择 β_k 使得算法有效，仍然是一个值得研究的问题．

从(3.4.11)和(3.4.18)我们看到，考虑一般形式的公式

$$d_{k+1}=-g_{k+1}+\tilde{\beta}_k d_k+\tilde{\gamma}_k d_{k-1} \tag{3.4.21}$$

和

$$d_{k+1}=-g_{k+1}+\hat{\beta}_k d_k+\hat{\gamma}_k y_k \tag{3.4.22}$$

是十分有意义的．

由于共轭梯度法具有二次终止性、内存需要量小，程序简单等优点，它将是无约束优化，特别是大规模问题的一个重要方法．对它的进一步研究是十分重要和必要的．

§3.5　一个一般性收敛定理

在本节中讨论线搜索方法的一个一般性收敛结果．先给出线搜索方法的一般形式．

算法 3.5.1

步1　给出 $x_1\in\mathbf{R}^n$，$0\leqslant\varepsilon<1$，$k=1$．

步2　如果 $\|\nabla f(x_k)\|_2\leqslant\varepsilon$，则停；

求出搜索方向 $d_k\in\mathbf{R}^n$，使其满足

$$d_k^T\nabla f(x_k)<0. \tag{3.5.1}$$

步3　利用某种线搜索条件求出步长 α_k；

步4　$x_{k+1}=x_k+\alpha_k d_k$；

$k:=k+1$，转步2．

算法 3.5.1 没有给出如何求搜索方向和计算搜索步长的具体方法，它是所有无约束优化的下降方向线搜索方法的概括形式．不难看出，算法 3.1.1 和 3.2.3 都是算法 3.5.1 的特殊形式．我们记 θ_k 为方向 d_k 和负梯度方向 $-\nabla f(x_k)$ 之间的夹角，于是有

$$\cos\theta_k=-\frac{d_k^T\nabla f(x_k)}{\|d_k\|_2\|\nabla f(x_k)\|_2}. \tag{3.5.2}$$

利用引理 2.5.3，很容易得到如下收敛性结果：

定理 3.5.2 设函数 $f(x)$ 连续可微，且 $\nabla f(x)$ 满足 Lipschitz 条件(2.5.13)，如果算法 3.5.1 产生的搜索方向 d_k 满足

$$\sum_{k=1}^{\infty}\cos^2\theta_k=+\infty, \tag{3.5.3}$$

而且线搜索条件(2.5.8)、(2.5.9)每次迭代都满足，则算法 3.5.1 产生的点列 x_k 必有限终止或者下面两极限

$$\liminf_{k\to\infty}\|\nabla f(x_k)\|_2=0, \tag{3.5.4}$$

$$\lim_{k\to\infty}f(x_k)=-\infty \tag{3.5.5}$$

必有一为真.

证明 从引理 2.5.3 可知

$$f(x_k)-f(x_{k+1})\geqslant\beta\cos^2\theta_k\|\nabla f(x_k)\|_2^2, \tag{3.5.6}$$

其中 $\beta=b_1(1-b_2)/M$ 是一与 k 无关的正常数. 如果算法不有限终止，则对一切 $k>1$，都有

$$\begin{aligned}f(x_1)-f(x_k)&=\sum_{i=1}^{k-1}[f(x_i)-f(x_{i+1})]\\&\geqslant\beta\min_{1\leqslant i\leqslant k}\|\nabla f(x_i)\|_2^2\sum_{i=1}^{k-1}\cos^2\theta_i.\end{aligned} \tag{3.5.7}$$

因为(3.5.3)成立，由不等式(3.5.7)即知极限(3.5.4)和(3.5.5)至少有一个成立. ∎

从定理 3.5.2 可知，只要夹角 θ_k 不趋于 $\pi/2$，或者趋于 $\pi/2$ 不太快(即(3.5.3)式满足)，则线搜索方法在一定意义下是全局收敛的. 不难看出，在定理 3.5.2 中将线搜索条件换成精确线性搜索后，该定理依然成立.

第 4 章

拟 牛 顿 法

拟牛顿法是一种基于逼近牛顿法的方法，它在每次迭代的搜索方向 d_k 满足

$$B_k d_k = -\nabla f(x_k), \tag{4.0.1}$$

其中 B_k 是一近似 $\nabla^2 f(x_k)$ 的矩阵. 如果 B_k 正定，拟牛顿法也称变尺度法. 拟牛顿法的基本思想是利用梯度差及步长构造矩阵满足拟牛顿方程. 拟牛顿法不必计算二阶导数，但在一定条件下是超线性收敛的，它是求解无约束优化所有利用一阶导数的方法中最有效的一类方法. 而从实际计算表明，BFGS 方法是拟牛顿法中最有效的一个方法. 关于拟牛顿法的研究很多，而且已十分成熟. 看来在拟牛顿法中找到比 BFGS 方法好得多的方法不大可能. 但如何使拟牛顿法能更快地修正矩阵，更适合于并行计算，仍有一些问题值得研究. 另外对拟牛顿法的理论研究，仍有一些重要问题，如第一个拟牛顿法 DFP 方法对于凸函数的收敛性，BFGS 方法在非凸目标函数下是否收敛等都还有待解决.

§4.1 牛 顿 法

解多维无约束优化问题的牛顿法是基于目标函数的二次泰勒展开. 因为在当前迭代点 x_k 附近，

$$f(x_k+d) \approx f(x_k) + d^T\nabla f(x_k) + \frac{1}{2}d^T\nabla^2 f(x_k)d. \tag{4.1.1}$$

所以，我们可考虑

$$\min \varphi_k(d) = d^T\nabla f(x_k) + \frac{1}{2}d^T\nabla^2 f(x_k)d. \tag{4.1.2}$$

假定 $\nabla^2 f(x_k)$ 是正定的，则知(4.1.2)的解是

$$d_k = -[\nabla^2 f(x_k)]^{-1}\nabla f(x_k). \tag{4.1.3}$$

牛顿法就是用(4.1.3)作为搜索方向的方法.

算法4.1.1

步1 $x_1 \in \mathbf{R}^n$ 给出，$0 \leqslant \varepsilon < 1$;

$k := 1$

步2 如果 $\|\nabla f(x_k)\|_2 \leqslant \varepsilon$，则停；

由(4.1.3)计算 d_k.

步3 利用线搜索求 $\alpha_k > 0$;

$x_{k+1} = x_k + \alpha_k d_k$; $k := k+1$; 转步2.

对于一致凸函数，我们有收敛性定理如下：

定理4.1.2 设 $f(x)$ 二次连续可微且一致凸，如果线搜索满足

$$f(x_k) - f(x_k + \alpha_k d_k) \geqslant \bar{\eta}\|g_k\|_2^2 \cos^2\langle d_k, -g_k\rangle, \tag{4.1.4}$$

其中 $g_k = \nabla f(x_k)$，$\bar{\eta} > 0$ 是一个与 k 无关的常数，则由算法4.1.1产生的点列 x_k 必满足

$$\lim \|\nabla f(x_k)\|_2 = 0, \tag{4.1.5}$$

且 x_k 收敛于 $f(x)$ 唯一的极小点.

证明 由于 $f(x)$ 一致凸，必存在常数 $\eta > 0$，使得

$$\sigma_n(\nabla^2 f(x)) \geqslant \eta \tag{4.1.6}$$

对一切 x 均成立. 由(4.1.4)可知 $f(x_k)$ 是严格单调下降的，故知点列 $\{x_k\}$ 必有界. 从而存在常数 $M > 0$，使得

$$\|\nabla^2 f(x_k)\|_2 \leqslant M \tag{4.1.7}$$

对一切 k 都成立. 由(4.1.3)、(4.1.6)和(4.1.7)可知

$$\cos^2(d_k, -g_k) \geqslant \frac{\eta}{M}. \tag{4.1.8}$$

因此得到

$$\infty > \sum_{k=1}^{\infty}[f(x_k) - f(x_{k+1})]$$

$$\geqslant \sum_{k=1}^{\infty} \eta\bar{\eta} M^{-1}\|g_k\|_2^2. \tag{4.1.9}$$

从(4.1.9)即知(4.1.5)成立. 由于 $f(x)$ 一致凸，故它只有一个稳

定点. 所以从(4.1.5)即知 x_k 必收敛于 $f(x)$ 唯一的极小点 x^*. ∎

假定 $f(x)$ 三次连续可微, 我们有

$$\nabla^2 f(x_k)[x_k+d_k-x^*] = -\nabla f(x_k)+\nabla^2 f(x_k)(x_k-x^*) = O(\|x_k-x^*\|_2^2). \tag{4.1.10}$$

如果点列 x_k 收敛于 $f(x)$ 的极小点 x^*, 且 $\nabla^2 f(x^*)$ 正定, 则从(4.1.10)可知对充分大的 k, 有

$$\|x_k+d_k-x^*\| = O(\|x_k-x^*\|_2^2). \tag{4.1.11}$$

对于精确线搜索, 我们有

$$d_k^T \nabla f(x_k+\alpha_k^* d_k) = 0. \tag{4.1.12}$$

于是存在 $t_k \in (0, \alpha_k^*)$, 使得

$$d_k^T[\nabla f(x_k)+\nabla^2 f(x_k+t_k d_k)\alpha_k^* d_k] = 0. \tag{4.1.13}$$

从而可知

$$\alpha_k^* = -\frac{g_k^T[\nabla^2 f(x_k)]^{-1} g_k}{g_k^T(\nabla^2 f(x_k))^{-1}\nabla^2 f(x_k+t_k d_k)(\nabla^2 f(x_k))^{-1} g_k}, \tag{4.1.14}$$

由(4.1.14)可得

$$\alpha_k^* = 1+O(\|x_k-x^*\|_2^2). \tag{4.1.15}$$

利用(4.1.15)和(4.1.11)即可推出:

$$\|x_k+\alpha_k^* d_k-x^*\| = O(\|x_k-x^*\|_2^2). \tag{4.1.16}$$

关系式(4.1.11)和(4.1.16)告诉我们步长为1的牛顿法和精确线搜索的牛顿法都是二次收敛的. 我们将其写成定理形式如下:

定理 4.1.3　设算法 4.1.1 所产生的点列 x_k 收敛于 x^*, $\nabla f(x^*)=0$, $\nabla^2 f(x^*)$ 正定. 如果 α_k 由精确线搜索求得, 或者对充分大的 k 有 $\alpha_k=1$, 则点列 x_k 二次收敛于 x^*, 即

$$\|x_{k+1}-x^*\| = O(\|x_k-x^*\|_2^2). \tag{4.1.17}$$

如果 $\nabla^2 f(x_k)$ 不正定, 则牛顿法的搜索方向(4.1.3)可能不是下降方向. 这时有必要对牛顿法进行修正. 我们把牛顿法的搜索方向称为牛顿步(Newton step), 记之为 d_k^N:

$$d_k^N = -[\nabla^2 f(x_k)]^{-1} g_k. \tag{4.1.18}$$

早期的修正牛顿步的方法是基于搜索方向与负梯度的夹角.

Goldstein 和 Price(1967)提出修正方法

$$d_k=\begin{cases}d_k^N, & \text{如果}\cos(d_k^N,\ -g_k)\geqslant\eta;\\ -g_k.\end{cases}\tag{4.1.19}$$

其中 $\eta>0$ 是一个预先给定的正常数，由于这样给定的搜索方向总有

$$\cos(d_k,\ -g_k)\geqslant\eta.\tag{4.1.20}$$

所以，算法的收敛性是可保证的．假定 x_k 收敛于 $f(x)$ 的极小点 x^*，且 $\nabla^2 f(x^*)$正定，则只要 $\eta<\sqrt{\sigma_n(\nabla^2 f(x^*))/\sigma_1(\nabla^2 f(x^*))}$，那么对充分大的 k 都有

$$d_k=d_k^N.\tag{4.1.21}$$

从而算法仍保持牛顿法的二次收敛性．

另一种修正牛顿法的方法是利用负曲率(negatire curvature)下降方向．这种方法最早由 Fiacco 和 McCormick(1968)提出．这方法每次迭代都试图对矩阵 $\nabla^2 f(x_k)$作 LDL^T 分解．如果所有的对角线元素 d_{ii} 都是正数，则 $\nabla^2 f(x_k)$ 是正定的，且这个 LDL^T 分解可用来计算牛顿步．如果存在 $d_{ii}<0$，则通过解方程

$$L^T d_k=\pm\sum_{\substack{i=1\\ d_{ii}<0}}^{n} e_i,\tag{4.1.22}$$

(其中 $e_i=(0,\ \cdots,\ 1,0,\cdots,0)^T$ 是第 i 个分量为 1，其他分量为 0 的向量)可求得一个负曲率方向 d_k，即

$$d_k^T\nabla^2 f(x_k)d_k<0.\tag{4.1.23}$$

我们可以适当地选取(4.1.22)右端的符号使得

$$d_k^T g_k\leqslant 0.\tag{4.1.24}$$

由于(4.1.23)和(4.1.24)，显然 d_k 是函数 $f(x)$在 x_k 点的下降方向．

Murray(1972) 提出一种对一般对称矩阵进行强制性 LL^T 分解，方法的实质是对 $\nabla^2 f(x_k)+D$ 进行 LL^T 分解．即在第 k 次迭代时，有

$$L_k L_k^T=\nabla^2 f(x_k)+D_k,\tag{4.1.25}$$

其中 D_k 是一对角阵．然后用公式

$$d_k = -L_k^{-T}L_k^{-1}g_k \tag{4.1.26}$$

产生搜索方向.

另一种与(4.1.25)十分类似的修正方法是令

$$d_k = -(\nabla^2 f(x_k) + \lambda_k I)^{-1} g_k, \tag{4.1.27}$$

其中 $\lambda_k \geqslant 0$ 是一参数. 当 $\nabla^2 f(x_k)$ 为正定的且条件数不太大时, 取 $\lambda_k = 0$, 否则, 取 λ_k 足够大, 使得矩阵 $(\nabla^2 f(x_k) + \lambda_k I)$ 的条件数不太大. 这个方法的思想可追溯到解非线性方程组的 Levenberg-Marquardt 方法(见 Levenberg, 1944; Marquardt, 1963). 不过, 这一方法的不便之处是对 λ_k 的确定并不容易. 人们发现, 在(4.1.27)中要求 $(\nabla^2 f(x_k) + \lambda_k I)$ 的条件数不太大, 基本上等价于要求 $\|d_k\|$ 不太大. 事实上, 假定 $\lambda_k \geqslant 0$, $(\nabla^2 f(x_k) + \lambda_k I)$ 正定, 且 d_k 由(4.2.27)给出, 则 d_k 是下面问题

$$\min \quad g_k^T d_k + \frac{1}{2} d^T \nabla^2 f(x_k) d, \tag{4.1.28}$$

$$\text{s. t} \quad \|d\|_2 \leqslant \Delta_k \tag{4.1.29}$$

的解, 其中 $\Delta_k = \|d_k\|$. (4.1.28)～(4.1.29)实质上就是一个信赖域方法的子问题, 人们发现, 调节 Δ_k 比调节 λ_k 更容易、更直观, 也更合理.

§ 4.2　拟牛顿法的导入

在每个迭代点 x_k 附近考虑 $f(x)$ 的二次逼近

$$f(x_k + d) \approx f(x_k) + d^T g_k + \frac{1}{2} d^T B_k d, \tag{4.2.1}$$

其中 $g_k = \nabla f(x_k)$, $B_k \in \mathbf{R}^{n \times n}$ 是对称矩阵. 如果 B_k 是非奇异的, 则(4.2.1)的右端函数的稳定点为

$$d_k = -B_k^{-1} g_k. \tag{4.2.2}$$

由于逼近关系(4.2.1), 我们有理由把 d_k 当做线搜索方向. 假定我们用某种方式已得到 B_k, 则利用搜索方向 d_k 进行一次线搜索得到 $x_{k+1} = x_k + \alpha_k d_k$. 现在的问题是如何确定 B_{k+1}? 显然, 我们希望 B_{k+1} 是 $\nabla^2 f(x_{k+1})$ 的某种近似. 由于

$$\nabla^2 f(x_{k+1}) s_k \approx y_k, \tag{4.2.3}$$

其中

$$s_k = x_{k+1} - x_k = \alpha_k d_k, \tag{4.2.4}$$

$$y_k = g_{k+1} - g_k = \nabla f(x_{k+1}) - \nabla f(x_k). \tag{4.2.5}$$

我们要求

$$B_{k+1} s_k = y_k. \tag{4.2.6}$$

(4.2.6) 被称为拟牛顿方程 (quasi-Newton equation) 或者拟牛顿条件. 从而可得到一般的拟牛顿算法如下:

算法 4.2.1

步 1 给出 $x_1 \in \mathbf{R}^n$, $0 \leqslant \varepsilon < 1$.
给出 B_1, $k := 1$;

步 2 如果 $\|g_k\|_2 \leqslant \varepsilon$, 则停.
计算 $d_k = -B_k^{-1} g_k$;
利用某种线搜索求 $\alpha_k > 0$;
$x_{k+1} = x_k + \alpha_k d_k$;

步 3 构造 B_{k+1} 使得(4.2.6)成立;
$k := k+1$, 转步 2.

我们记 $H_k = B_k^{-1}$, 则(4.2.6)可写成

$$H_{k+1} y_k = s_k. \tag{4.2.7}$$

关于拟牛顿法, 目前记号比较混乱. 早期人们都用 $\delta_k = x_{k+1} - x_k$, $\gamma_k = g_{k+1} - g_k$, 由于国际上通用小写希腊字母表示标量, 现在逐渐改用(4.2.4)和 (4.2.5). 我们用 B_k 来表示近似的 $\nabla^2 f(x_k)$, H_k 表示 B_k^{-1} 是传统的记号(见 Fletcher, 1987). 有的人根据 H 是海色阵 (Hessian) 英文单词的第一个字母而用 H_k 来表示近似的 $\nabla^2 f(x_k)$ (见 Dennis 和 Schnabel, 1983). 这两种表达方式下的 H_k 正好互为逆矩阵. 最近, 为了避免混淆, Davidon, Mifflin 和 Nazareth(1991) 提出用 M_k 来表示近似的 $\nabla^2 f(x_k)$ 和 $W_k^{-1} = M_k$. 他们的主要理由是: 第一, M 和 W 都是对称字母; 第二, M 倒过来正好是 W. 但现在还不知他们这种统一记号的努力是否能达到目的.

假定 B_k 正定，不难证明由(4.2.2)定义的 d_k 是函数 $f(x)$ 在 x_k 处在范数 $\|x\|=x^TB_kx$ 意义下的最速下降方向. 所以，如果所有的 B_k 都是正定的，则算法 2.4.1 也称为变尺度(variable metric)方法.

§ 4.3 几个重要的拟牛顿法

第一个拟牛顿法由 Davidon(1959) 提出，后经 Fletcher 和 Powell(1963)修改及整理成以下给出的形式，故称为 DFP 方法. 它的修正公式为

$$B_{k+1}=B_k-\frac{B_ks_ky_k^T+y_ks_k^TB_k}{s_k^Ty_k}+\left(1+\frac{s_k^TB_ks_k}{s_k^Ty_k}\right)\frac{y_ky_k^T}{s_k^Ty_k}. \tag{4.3.1}$$

不难验证，从(4.2.1)可推出 $H=B^{-1}$ 的修正公式:

$$H_{k+1}=H_k-\frac{H_ky_ky_k^TH_k}{y_k^TH_ky_k}+\frac{s_ks_k^T}{s_k^Ty_k}. \tag{4.3.2}$$

由于精确线搜索($s_k^T\nabla f(x_{k+1})=0$)和非精确线搜索(2.5.9) 都满足

$$s_k^Ty_k>0. \tag{4.3.3}$$

所以，我们可假定(4.3.3) 对一切 k 均满足. 下面的结果是显然的.

定理 4.3.1 设 B_k 为对称正定阵，且 (4.3.3) 成立，则由(4.3.1)定义的 B_{k+1} 也对称正定，满足拟牛顿方程 (4.2.6)，而且

$$\det(B_{k+1})=\frac{y_k^TH_ky_k}{s_k^Ty_k}\det(B_k). \tag{4.3.4}$$

证明 我们可将(4.3.1)写成

$$B_{k+1}=\left(I-\frac{y_ks_k^T}{s_k^Ty_k}\right)B_k\left(I-\frac{s_ky_k^T}{s_k^Ty_k}\right)+\frac{y_ky_k^T}{s_k^Ty_k}. \tag{4.3.5}$$

则显然可见 B_{k+1} 对称正定且满足拟牛顿方程. 从(4.3.5)还可推出 (记 $\tilde{s}_k=B_k^{\frac{1}{2}}s_k$, $\tilde{y}_k=B_k^{-\frac{1}{2}}y_k$):

$$B_k^{-\frac{1}{2}} B_{k+1} B_k^{-\frac{1}{2}} = \left(I - \frac{\tilde{y}_k \tilde{s}_k^T}{\tilde{s}_k^T \tilde{y}_k}\right)\left(I - \frac{\tilde{s}_k \tilde{y}_k^T}{\tilde{s}_k \tilde{y}_k}\right) + \frac{\tilde{y}_k \tilde{y}_k^T}{\tilde{s}_k^T \tilde{y}_k}. \tag{4.3.6}$$

不难看出(4.3.6)右端的矩阵有 $n-2$ 个特征值为 1；另两个特征值满足方程

$$t^2 - \left(\frac{\|\tilde{s}_k\|_2^2 \|\tilde{y}_k\|_2^2}{(\tilde{s}_k^T \tilde{y}_k)^2} + \frac{\|\tilde{y}_k\|_2^2}{\tilde{s}_k^T \tilde{y}_k}\right) t + \frac{\|\tilde{y}_k\|_2^2}{\tilde{s}_k \tilde{y}_k} = 0. \tag{4.3.7}$$

从(4.3.7)和(4.3.6)即知

$$\det(B_k^{-\frac{1}{2}} B_{k+1} B_k^{-\frac{1}{2}}) = \frac{\|\tilde{y}_k\|_2^2}{\tilde{s}_k^T \tilde{y}_k} = \frac{y_k^T H_k y_k}{s_k^T y_k}. \tag{4.3.8}$$

(4.3.8)显然与(4.3.4)等价. ▌

另一个著名的拟牛顿法由 Broyden(1970)、Fletcher(1970)、Goldfarb(1970) 及 Shanno(1970) 独立提出的，所以常被称为 BFGS 方法，它的矩阵修正公式为

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{4.3.9}$$

$$H_{k+1} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{y_k^T s_k} + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k}. \tag{4.3.10}$$

比较(4.3.1)、(4.3.2)和(4.3.9)、(4.3.10)，不难发现 BFGS 方法的修正公式与 DFP 方法的修正公式在形式上一模一样. 更精确地说，如果在(4.3.1)、(4.3.2)式中令 B_k 与 H_k 对换，s_k 与 y_k 对换就得到了(4.3.9)和(4.3.10). 正由于这种关系，(4.3.9)与(4.3.10)称为(4.3.1)与(4.3.2)的对偶公式，BFGS 方法和 DFP 方法称为互为对偶方法. 由于对偶性和定理 4.3.1，从而显然有以下定理.

定理 4.3.2 设 B_k 对称正定且(4.3.3)成立，则由(4.3.9)定义的 B_{k+1} 也对称正定且满足拟牛顿方程(4.2.6)，而且

$$\det(B_{k+1}) = \frac{s_k^T y_k}{s_k^T B_k s_k} \det(B_k). \tag{4.3.11}$$

对于任何两个矩阵 $\bar{B}$、$\hat{B}$，如果它们都满足拟牛顿公式

(4.2.6), 即

$$\bar{B}s_k = y_k,\ \hat{B}s_k = y_k. \tag{4.3.12}$$

则对任何 $\theta \in \mathbf{R}^1$, 矩阵

$$\bar{B} + \theta(\hat{B} - \bar{B}) \tag{4.3.13}$$

必满足拟牛顿公式 (4.2.6). Broyden 利用 BFGS 和 DFP 公式构造出一族拟牛顿修正公式:

$$B_{k+1}(\theta) = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \theta w_k w_k^T, \tag{4.3.14}$$

其中

$$w_k = \sqrt{s_k^T B_k s_k}\left(\frac{y_k}{s_k^T y_k} - \frac{B_k s_k}{s_k^T B_k s_k}\right), \tag{4.3.15}$$

$\theta \in \mathbf{R}^1$ 是一参数. 在(4.3.14)中令 $\theta = 0$, 就得到了 BFGS 修正公式; 令 $\theta = 1$, 就得到了 DFP 修正公式. 从(4.3.14)的形式可知 $\det(B(\theta))$ 必为 θ 的线性函数, 再利用(4.3.11)和(4.3.4)即知

$$\det(B_k(\theta)) = \det(B_k)[1 + \theta(\beta_k\gamma_k - 1)]/\gamma_k, \tag{4.3.16}$$

其中

$$\beta_k = \frac{y_k^T H_k y_k}{s_k^T y_k},\ \gamma_k = \frac{s_k^T B_k s_k}{s_k^T y_k}. \tag{4.3.17}$$

如果 B_k 正定, 则 $\beta_k\gamma_k \geqslant 1$, 且

$$\beta_k\gamma_k - 1 = w_k^T H_k w_k, \tag{4.3.18}$$

由(4.3.18)可知, 如果 $\beta_k\gamma_k - 1 = 0$, 则 $w_k = 0$, 从而 Broyden 族中的所有方法都一样. 假定 B_k 非奇异且 $\beta_k\gamma_k - 1 \neq 0$, 由(4.3.16)可知, 只要

$$\theta \neq 1/(1 - \beta_k\gamma_k), \tag{4.3.19}$$

则 $B_{k+1}(\theta)$ 可逆. 我们可把 $B_{k+1}(\theta)$ 的逆写成

$$H_{k+1}(\phi) = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{s_k s_k^T}{s_k^T y_k} + \phi v_k v_k^T, \tag{4.3.20}$$

其中

$$v_k = \sqrt{y_k^T H_k y_k}\left(\frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k}\right), \tag{4.3.21}$$

且

$$\phi = \phi(\theta) = (1 - \theta)/(1 + \theta(\beta_k\gamma_k - 1)). \tag{4.3.22}$$

假定 B_k 正定且 $\beta_k\gamma_k-1>0$, 则对任何 $\theta>\frac{1}{1-\beta_k\gamma_k}$, $B_{k+1}(\theta)$也是对称正定阵. 所以我们称 $\theta>\frac{1}{1-\beta_k\gamma_k}$ 的 Broyden 子族为 Broyden 正族. 不难发现, 由 (4.3.22) 定义的函数 $\phi(\theta)$ 是区间 $\left(\frac{1}{1-\beta_k\gamma_k},\ +\infty\right)$上的严格单调下降函数, 且是该区间上的一对一影射. 所有的 Broyden 正族也是指对所有的 $\phi>\frac{1}{1-\beta_k\gamma_k}$ 的修正公式. 另一个有趣性质是 $\phi(\theta)$也是[0, 1]上的一对一影射. 也就是说, 如果矩阵 B 是(4.3.1)和(4.3.9)中 B_{k+1} 的凸组合, 则 B 的逆矩阵必是 (4.3.2) 和 (4.3.10) 中 H_{k+1} 的凸组合. 所以 Broyden 族(4.3.14)当 $\theta\in[0,\ 1]$时称为 Broyden 凸族.

在 Broyden 凸族中令

$$\theta=\frac{s_k^T y_k}{s_k^T y_k+s_k^T B_k s_k},\quad \phi=\frac{s_k^T y_k}{s_k^T y_k+y_k^T H_k y_k}. \tag{4.3.23}$$

则得到修正公式

$$B_{k+1}=B_k-\frac{(B_k s_k+y_k)(B_k s_k+y_k)^T}{s_k^T y_k+s_k^T B_k s_k}+2\frac{y_k y_k^T}{s_k^T y_k}, \tag{4.3.24}$$

$$H_{k+1}=H_k-\frac{(H_k y_k+s_k)(H_k y_k+s_k)^T}{s_k^T y_k+y_k^T H_k y_k}+2\frac{s_k s_k^T}{s_k^T y_k}. \tag{4.3.25}$$

它最先是由 Hoshino(1972)导出的, 所以被称为 Hoshino 修正公式. 在(4.3.24)与(4.3.25)中把 B 和 H 对换, s_k 和 y_k 对换, 得到的公式仍然是 (4.3.24) 与 (4.3.25). 由于这一性质, 我们称(4.3.24)与(4.3.25) 是自对偶的 (self-dual). Broyden 凸族的另一个自对偶的方法可由

$$\theta=\phi=\frac{1}{1+\sqrt{\beta_k\gamma_k}} \tag{4.3.26}$$

得到:

$$B_{k+1}=B_k-\frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}+\frac{y_k y_k^T}{s_k^T y_k}+\frac{1}{1+\sqrt{\beta_k\gamma_k}}w_k w_k^T, \tag{4.3.27}$$

$$H_{k+1}=H_k-\frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}+\frac{s_k s_k^T}{s_k^T y_k}+\frac{1}{1+\sqrt{\beta_k\gamma_k}}\,v_k v_k^T. \tag{4.3.28}$$

这个修正公式最早是由谢元富(1989)推出的，后来 Yuan(1990b) 在研究 Broyden 族中所有可能对偶公式时独立地导出. 修正公式 (4.3.27) 与 (4.3.28) 是 Broyden 正族中唯一满足 $\theta=\phi$ 的公式，它可以理解为 DFP 方法和 BFGS 方法的几何平均值 (见 Yuan, 1990b).

以上提到的修正公式在一般情况下均有 $B_{k+1}-B_k$ 的秩为 2，故这些修正公式均称为秩为 2(Rank 2) 修正公式. 唯一的一个对称秩为 1 的修正公式是

$$B_{k+1}=B_k+\frac{(y_k-B_k s_k)(y_k-B_k s_k)^T}{(y_k-B_k s_k)^T s_k}. \tag{4.3.29}$$

它的逆矩阵的修正公式为

$$H_{k+1}=H_k+\frac{(s_k-H_k y_k)(s_k-H_k y_k)^T}{(s_k-H_k y_k)^T y_k}. \tag{4.3.30}$$

对称秩为 1 的算法由 Broyden(1967)、Davidon(1968)、Fiacco 和 McCormick(1968)、Murtagh 和 Sargent(1969)以及 Wolfe(1968) 独立提出的. 对称秩为 1 的修正公式也属于 Broyden 族，它相应的参数是

$$\theta=\frac{s_k^T y_k}{s_k^T y_k-s_k^T B_k s_k},\quad \phi=\frac{s_k^T y_k}{s_k^T y_k-y_k^T H_k y_k}. \tag{4.3.31}$$

对称秩为 1 的公式也是一个自对偶公式. 从(4.3.29)与(4.3.30)可以看出如果 $\beta_k=1$，则 B_{k+1} 奇异；如果 $\gamma_k=1$，则 H_{k+1} 奇异. 而且有

定理 4.3.3 设 B_k 对称正定且(4.3.3)成立，则(4.3.29) 定义的 B_{k+1} 也对称正定当且仅当

$$\min\{\beta_k,\ \gamma_k\}<1. \ ▮ \tag{4.3.32}$$

证明 如果(4.3.32)成立，则从(4.3.31)即知

$$\max\{\theta,\ \phi\}>0>\frac{1}{1-\beta_k\gamma_k}. \tag{4.3.33}$$

从而可知 B_{k+1} 必对称正定. 如果 B_{k+1} 由(4.3.29)定义对称正定, 则知由(4.3.31)定义的 θ、ϕ 必满足:

$$\theta > \frac{1}{1-\beta_k\gamma_k}, \quad \phi > \frac{1}{1-\beta_k\gamma_k}. \tag{4.3.64}$$

由(4.3.34)、(4.3.31)以及 β_k、γ_k 的定义即知(4.3.32)必成立. ∎

一族包含三个参数的修正公式由 Huang (1970) 给出的. 它的形式是

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \rho_k \frac{s_k s_k^T}{s_k^T y_k} + (\tau_k s_k + \xi_k H_k y_k) v_k^T. \tag{4.3.35}$$

其中 ρ_k、τ_k、ξ_k 是三个参数, v_k 由(4.3.21)定义. Huang 族的修正公式满足

$$H_{k+1} y_k = \rho_k s_k, \tag{4.3.36}$$

除非 $\rho_k = 1$, Huang 族中的修正公式不满足拟牛顿方程. 如果要求 Huang 族中的修正矩阵对称, 则(4.3.35)必具有形式

$$H_{k+1} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \rho_k \frac{s_k s_k^T}{s_k^T y_k} + \phi_k v_k v_k^T. \tag{4.3.37}$$

从上式不难看出 Broyden 族就是 Huang 族中所有对称且满足拟牛顿方程的所有修正公式. 修正公式(4.3.35)的逆矩阵修正公式具有形式:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{1}{\rho_k} \frac{y_k y_k^T}{s_k^T y_k} + (\tau'_k y_k + \xi'_k B_k s_k) w_k^T. \tag{4.3.38}$$

另一个形式简单的秩为 1 的修正公式最先由 McCormick 提出(见 Pearson, 1969), 它具有形式:

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k) u_k^T}{u_k^T y_k}. \tag{4.3.39}$$

其中 u_k 是一满足 $u_k^T y_k \neq 0$ 的向量. (4.3.39)的逆矩阵形式为

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) u_k^T B_k}{u_k^T B_k s_k}. \tag{4.3.40}$$

在公式 (4.3.39) 与 (4.3.40) 中令 $u_k = H_k s_k$, 就得到了 Broyden

(1965)非对称秩为1的修正公式:

$$B_{k+1}=B_k+\frac{(y_k-B_ks_k)s_k^T}{s_k^Ts_k},\tag{4.3.41}$$

$$H_{k+1}=H_k+\frac{(s_k-H_ky_k)s_k^TH_k}{s_k^TH_ky_k}.\tag{4.3.42}$$

用非对称阵去近似 $\nabla^2f(x_k)$ 看来不太合理，但在解非线性方程组的拟牛顿法中，修正矩阵是用来逼近非线性方程组的雅可比(Jacobi)阵，所以非对称秩为1的修正公式确是十分合理的，而且在那种情形(即解方程组)，如果采用对称修正公式，则每个近似矩阵均对称，这倒反而不妥，因为在一般情况下，雅可比阵是非对称的.

如果我们将 Broyden 非对称秩为1的修正公式对称化，则可得到下列公式:

$$B_{k+1}=B_k+\frac{(y_k-B_ks_k)s_k^T+s_k(y_k-B_ks_k)^T}{s_k^Ts_k}-\frac{s_k^T(y_k-B_ks_k)s_ks_k^T}{(s^Ts_k)^2}.\tag{4.3.43}$$

它由 Powell(1970)得到，(4.3.43)的逆矩阵的修正公式十分复杂，故一般只用修正 B_k 而不计算 H_k. 由于 Powell 的修正公式是下列数列的极限. 即:

$$B_k^{(0)}=B_k,\tag{4.3.44}$$

$$\overline{B}_k^{(i)}=B_k^{(i)}+\frac{(y_k-B_k^{(i)}s_k)s_k}{s_k^Ts_k},\tag{4.3.45}$$

$$B_k^{(i+1)}=[\overline{B}_k^{(i)}+(\overline{B}_k^{(i)})^T]/2.\tag{4.3.46}$$

则由(4.3.43)定义的 B_{k+1} 是 $B_k^{(i)}$ 当 i 趋于无穷时的极限. 故修正公式可理解为对称化的 Broyden 秩为1的公式. 所以人们常称(4.3.43)为 Powell 的对称 Broyden 公式，或简称 PSB(Powell's Symmetric Broyden) 公式. 值得提出的是 PSB 公式并不属于 Huang 族，从而也不属于 Broyden 族.

Davidon(1975)提出修正公式

$$H_{k+1}=H_k-\frac{H_k\bar{y}_k\bar{Y}_k^TH_k}{\bar{Y}_k^TH_k\bar{Y}_k}+\frac{\bar{s}_k\bar{s}_k^T}{\bar{s}_k^T\bar{y}_k}+\phi\bar{v}_k\bar{v}_k^T \tag{4.3.47}$$

其中 ϕ 是参数，且

$$\bar{v}_k=\sqrt{\bar{y}_k^TH_k\bar{y}_k}\left(\frac{\bar{s}_k}{\bar{s}_k^T\bar{y}_k}-\frac{H_k\bar{y}_k}{\bar{y}_k^TH_k\bar{y}_k}\right). \tag{4.3.48}$$

$\bar{s}_k$、$\bar{y}_k\in \mathbf{R}^n$ 由某种方式产生. 显然，当 $\bar{s}_k=s_k$、$\bar{y}_k=y_k$ 时，则 Davidon 族(4.3.47)就是 Broyden 族.

最近，Yuan(1990b)导出了 Broyden 正族中的另一个自对偶修正公式，即

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_ky_k^T}{s_k^Ty_k}+\frac{\beta_k-1}{\beta_k\gamma_k-1}w_kw_k^T, \tag{4.3.49}$$

$$H_{k+1}=H_k-\frac{H_ky_ky_k^TH_k}{y_k^TH_ky_k}+\frac{s_ks_k^T}{s_k^Ty_k}+\frac{\gamma_k-1}{\beta_k\gamma_k-1}v_kv_k^T. \tag{4.3.50}$$

由于 β_k-1 和 γ_k-1 两者必有一为非负，故知修正公式(4.3.49)与(4.3.50)是属于 Broyden 正族. 当 $\beta_k\ll\gamma_k$ 时，这个公式十分靠近 BFGS 公式；而当 $\beta_k\gg\gamma_k$ 时，它十分靠近 DFP 公式. 由于 BFGS 公式在 $\beta_k\ll\gamma_k$ 时比 DFP 有效，而 DFP 在 $\beta_k\gg\gamma_k$ 时比 BFGS 有效(见 Powell,1986)，所以修正公式(4.3.49)与(4.3.50)可理解为介于 BFGS 和 DFP 之间的一个自动调节公式.

另一类修正公式具有下列形式:

$$B_{k+1}=\rho\left(B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\theta w_kw_k^T\right)+\frac{y_ky_k^T}{s_k^Ty_k}, \tag{4.3.51}$$

$$H_{k+1}=\frac{1}{\rho}\left(H_k-\frac{H_ky_ky_k^TH_k}{y_k^TH_ky_k}+\phi v_kv_k^T\right)+\frac{s_ks_k^T}{s_k^Ty_k}. \tag{4.3.52}$$

这类修正公式被称为自调节(self-scaling)公式，它最早由 Oren (1972)提出的，从(4.3.51)可看出，自调节公式可理解为先将 B_k 乘上 ρ，然后用 Broyden 族的公式得到 B_{k+1}；也可以理解为先利

用 Huang 族中的对称公式(4.3.37)求得 B_{k+1}, 然后将 B_{k+1} 乘以 ρ. 自调节方法中的一个特殊公式是自调节 BFGS:

$$B_{k+1}=\frac{s_k^T y_k}{s_k^T B_k s_k}\left(B_k-\frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}\right)+\frac{y_k y_k^T}{s_k^T y_k}. \qquad (4.3.53)$$

它是由 Shanno 和 Phua(1978)给出的. 显然, 自调节 BFGS 方法也是一个变尺度法. 关于其他自调节方法的介绍可参阅 Oren (1982)以及 Oren 和 Luenberger(1974).

§4.4 不变性和二次终止性

拟牛顿方法的一个重要性质是不变性(invariance), 即它在经过变量线性变换后保持不变. 设 A 是一非奇异阵, $a\in\mathbf{R}^n$, 考虑线性变换

$$\tilde{x}=Ax+a, \qquad (4.4.1)$$

则目标函数变成

$$\tilde{f}(\tilde{x})=f(A^{-1}(\tilde{x}-a))=f(x). \qquad (4.4.2)$$

不难求得

$$\nabla\tilde{f}(\tilde{x})=A^{-T}\nabla f(x). \qquad (4.4.3)$$

利用(4.4.1)和(4.4.3)可得到以下引理:

引理 4.4.1 设 $\tilde{x}_k=Ax_k+a$, 且

$$\tilde{H}_k=AH_kA^T, \qquad (4.4.4)$$

则有

$$\tilde{d}_k=-\tilde{H}_k\tilde{g}_k=Ad_k, \qquad (4.4.5)$$

$$\tilde{d}_k^T\tilde{g}_k=d_k^Tg_k. \qquad (4.4.6)$$

如果 $\tilde{\alpha}_k=\alpha_k$, 则

$$\tilde{x}_{k+1}=Ax_{k+1}+a, \qquad (4.4.7)$$

且

$$\tilde{y}_k=A^{-T}y_k. \qquad (4.4.8)$$

证明 (4.4.5) 可由 (4.4.3) 和 (4.4.4) 直接推出. 于是由(4.4.5)和(4.4.3)即知(4.4.6)成立. 如果 $\tilde{\alpha}_k=\alpha_k$, 则有

$$\begin{aligned}\tilde{x}_{k+1}&=\tilde{x}_k+\tilde{\alpha}_k\tilde{d}_k\\&=Ax_k+a+\alpha_kAd_k\\&=A[x_k+\alpha_kd_k]+a\\&=Ax_{k+1}+a.\end{aligned}\tag{4.4.9}$$

即知(4.4.7)成立．由(4.4.7)知 $\tilde{g}_{k+1}=A^{-T}g_{k+1}$，从而(4.4.8)必成立． ▌

上面的引理说明了只要拟牛顿矩阵满足一定关系，经过线性变换后的拟牛顿法和原来的方法产生一样的方向．所以只要每次迭代都有 $\tilde{\alpha}_k=\alpha_k$ 和关系式(4.4.4)成立，则对所有的 k，均有

$$\tilde{x}_k=Ax_k+a.\tag{4.4.10}$$

从而可知，只要(4.4.4)式成立，且 α_k 的选取不受影响，则线性变换不会改变拟牛顿法产生的点列$\{x_k\}$．显然，在精确搜索下 $\tilde{\alpha}_k=\alpha_k$．由于(4.4.6)以及

$$\tilde{f}(\tilde{x}_k)=f(x_k),\tag{4.4.11}$$

$$\tilde{f}(\tilde{x}_k+\alpha\tilde{d}_k)=f(x_k+\alpha d_k),\tag{4.4.12}$$

$$\tilde{d}_k^T\nabla f(\tilde{x}_k+\alpha\tilde{d}_k)=d_k^T\nabla f(x_k+\alpha d_k).\tag{4.4.13}$$

我们知道如果非精确线搜索是基于 $f(x_k)$、$d_k^T\nabla f(x_k)$、$f(x_k+\alpha d_k)$、$d_k^T\nabla f(x_k+\alpha d_k)$，则有 $\tilde{\alpha}_k=\alpha_k$．不难发现非精确线搜索(2.5.8)与(2.5.9)将有 $\tilde{\alpha}_k=\alpha_k$．假定 $\tilde{\alpha}_k=\alpha_k$，且 H_{k+1} 和 $\tilde{H}_{k+1}$ 分别由修正公式

$$H_{k+1}=H_k+(\tau s_k+\xi H_ky_k)(\tau s_k+\xi H_ky_k)^T\tag{4.4.14}$$

$$\tilde{H}_{k+1}=\tilde{H}_k+(\tau\tilde{s}_k+\xi\tilde{H}_k\tilde{y}_k)(\tau\tilde{s}_k+\xi\tilde{H}_k\tilde{y}_k)^T\tag{4.4.15}$$

得到，则有

$$\tilde{H}_{k+1}=AH_kA^T.\tag{4.4.16}$$

由于 $s_k^Ty_k$、$s_k^TB_ks_k$、$y_k^TH_ky_k$ 都是不变量，于是有

定理 4.4.2 假定 $\tilde{x}_1=Ax_1+a$，$\tilde{H}_1=AH_1A^T$，如果 $\tilde{\alpha}_k=\alpha_k$ 且 H_k 和 $\tilde{H}_k$ 由 Huang 族中同一修正公式得到，则对一切 k，都有

$$\tilde{x}_k=Ax_k+a.\tag{4.4.17}$$

定理 4.4.2 说明只要修正公式属于 Huang 族，则线性变换并不会改变拟牛顿法产生的点列．于是对于一个非常坏的条件

(illcondition) 的问题，即矩阵 $\nabla^2 f(x)$ 条件数很坏，我们可利用线性变换(4.4.1)，使得等价问题

$$\min \tilde{f}(\tilde{x}) \tag{4.4.18}$$

并不是坏条件问题．由于拟牛顿法的不变性，我们可知，这类方法对坏条件问题并不太敏感，而最速下降法显然没有不变性，故当 $\nabla^2 f(x)$ 的条件数很大时，收敛相当慢．

由于拟牛顿法的不变性，故在分析算法时，总可假定 $H_1=I$．下面的结果正是利用这一假设而得到的．

定理 4.4.3　假定目标函数 $f(x)$ 由(3.2.1)给出，α_k 由精确线搜索给出，$H_1=I$, $H_k(k>1)$ 由 Huang 族中的修正公式(4.3.37)给出，则对任何 k，都有

$$g_{k+1}^T s_i=0,\ i=1,\ \cdots,\ k; \tag{4.4.19}$$

$$H_{k+1} y_i=\rho_i s_i,\ i=1,\ \cdots,\ k; \tag{4.4.20}$$

$$s_i^T H s_j=0,\ i\neq j;\ i、j=1,\ \cdots,\ k+1. \tag{4.4.21}$$

证明　当 $k=1$ 时，(4.4.19)～(4.4.21)显然成立．假定对某 $k\geqslant 1$，(4.4.19)～(4.4.21)成立．我们证明这些关系式对 $k+1$ 也成立．

显然，$g_{k+2}^T s_{k+1}=0$，对于 $i=1,\ \cdots,\ k$，我们有

$$g_{k+2}^T s_i=y_{k+1}^T s_i+g_{k+1}^T s_i=d_{k+1}^T H s_i=0. \tag{4.4.22}$$

所以(4.4.19)对 $k+1$ 也成立．我们由修正公式知：

$$H_{k+2} y_{k+1}=\rho_{k+1} s_{k+1}. \tag{4.4.23}$$

对于 $i=1, 2, \cdots, k$，必存在向量 a_k、b_k，使得

$$\begin{aligned} H_{k+2} y_i &= (H_{k+2} H_{k+1}^{-1}) H_{k+1} y_i \\ &= (I+a_k y_{k+1}^T+b_k s_{k+1} H_{k+1}^{-1})\rho_i s_i \\ &= \rho_i s_i+a_k (g_{k+2}-g_{k+1})^T s_i+b_k s_{k+1}^T H s_i \\ &= \rho_i s_i. \end{aligned} \tag{4.4.24}$$

现在我们只需证明(4.4.21)对 $k+1$ 也成立．不难看出，我们只证 $i=k+2$ 就够了．利用已证明的(4.4.22)和(4.4.24)，有

$$\begin{aligned} d_{k+2}^T H s_i &= d_{k+2}^T y_i = -g_{k+2}^T H_{k+2} y_i \\ &= -\rho_i g_{k+2}^T s_i = 0. \end{aligned} \tag{4.4.25}$$

由归纳法即知定理成立. ∎

从(4.4.21)可知,必存在最小的整数 $m\leqslant n$,使得

$$s_{m+1}=0.$$

由于 $s_1, \cdots, s_m$ 相互 H-共轭,且有

$$\operatorname{span}\{g_1, g_2, \cdots, g_m\}=\operatorname{span}\{s_1, \cdots, s_m\}, \tag{4.4.26}$$

故知

$$g_{m+1}^T y_i=0,\ i=1, 2, \cdots, m-1. \tag{4.4.27}$$

由(4.4.27)、(4.4.19)以及 $H_1=I$,即知

$$g_{m+1}^T H_i g_{m+1}=\|g_{m+1}\|_2^2 \tag{4.4.28}$$

对一切 $i=1, \cdots, m$ 均成立. 于是

$$\begin{aligned} g_{m+1}^T H_{m+1} g_{m+1} &= g_{m+1}^T H_m g_{m+1}-\frac{[g_{m+1}^T H_m y_m]^2}{y_m^T H_m y_m} \\ &\quad +\phi_m (v_m^T g_{m+1})^2 \\ &= \|g_{m+1}\|_2^2 \frac{g_m^T H_m g_m}{y_m^T H_m y_m}+\phi_m \frac{\|g_{m+1}\|_2^4}{y_m^T H_m y_m}. \end{aligned} \tag{4.4.29}$$

因为 $s_{m+1}=0$,故必有 $g_{m+1}^T H_{m+1} g_{m+1}=0$. 如果 $\|g_{m+1}\|_2\neq 0$,则 $\phi_m=-g_m^T H_m g_m/\|g_{m+1}\|_2^2$. 所以,只要假定对一切 k,有

$$\phi_k\neq -g_k^T H_k g_k/\|g_{k+1}\|_2^2, \tag{4.4.30}$$

则必有

$$\|g_{m+1}\|=0. \tag{4.4.31}$$

也就是说,如果(4.4.30)得到满足,则 Huang 族中的任何方法在精确线搜索下具有二次终止性. 我们将其写成定理形式如下:

定理 4.4.4 在定理 4.4.3 的假定下,如果对一切 k(4.4.30)得到满足,则算法必有限终止,即存在 $m\leqslant n$,使得(4.4.31)成立,而且由算法产生的点列$\{x_k\}$和共轭梯度法产生的点列一样.

证明 (4.4.31)已经被证明了. 由于 $s_1, \cdots, s_m$ 相互 H-共轭,以及线搜索是由精确搜索得到的,不难证明 x_i 是函数 $f(x)$在

$$x_1+\operatorname{span}\{s_1, \cdots, s_{i-1}\}=x_1+\operatorname{span}\{g_1, \cdots, g_{i-1}\} \tag{4.4.32}$$

子空间上的最小值，从而必和共轭梯度法的 x_i 为同一点. ∎

定理 4.4.4 的一个直接推论如下：

推论 4.4.5 假定目标函数 $f(x)$ 由(3.2.1)给出，α_k 由精确线搜索给出，$H_1=I$，$H_k(k>1)$ 由 Broyden 正族的任何修正公式给出，则拟牛顿法 4.2.1 必有限终止. 即存在 $m\leqslant n$，使得(4.4.31)成立.

证明 由于矩阵 H_k 是由 Broyden 正族中的修正公式给出的，所有对一切 k，都有 H_k 对称正定. 于是只要 $\|g_{k+1}\|\neq 0$，则必有

$$g_{k+1}^T H_{k+1} g_{k+1}\neq 0. \tag{4.4.33}$$

从而可知(4.4.30)必满足，所以推论成立. ∎

由于拟牛顿法的不变性，将推论 4.4.5 中的 $H_1=I$ 改为 H_1 为任何对称正定阵，推论仍然成立. 拟牛顿法的二次终止性是基于精确线搜索，但对于对称秩为 1 的算法，这一条件并不必要.

定理 4.4.6 设 H_1 是任意给定的对称阵，并且对于任何 k，$(s_k-H_k y_k)^T y_k\neq 0$. H_{k+1} 是由对称秩为 1 的公式(4.3.30)给出的，如果 $s_1, s_2, \cdots, s_n$ 线性无关，则必有 $H_{n+1}=H^{-1}$.

证明 我们证明

$$H_{i+1} y_j = s_j (j=1, \cdots, i) \tag{4.4.34}$$

对一切 $i=1, \cdots, n$ 均成立. 显然(4.4.34)对 $i=1$ 成立. 假定(4.4.34)对某 $i\geqslant 1$ 成立，则有

$$(s_{i+1}-H_{i+1}y_{i+1})^T y_j = S_{i+1}^T y_j - y_{i+1}^T H_{i+1}^T y_j = 0 \tag{4.4.35}$$

对一切 $j=1, \cdots, i$ 都成立. 于是

$$\begin{aligned} H_{i+2} y_j &= H_{i+1} y_j + \frac{(s_{i+1}-H_{i+1}y_{i+1})(s_{i+1}-H_{i+1}y_{i+1})^T y_j}{(s_{i+1}-H_{i+1}y_{i+1})^T y_{i+1}} \\ &= H_{i+1} y_j = s_j, \end{aligned} \tag{4.4.36}$$

对一切 $j=1, \cdots, i$ 都成立. 由修正公式(4.3.30)还可知

$$H_{i+2} y_{i+1} = s_{i+1}. \tag{4.4.37}$$

由(4.4.36)和(4.4.37)即知(4.4.34)对 $i+1$ 也成立，由归纳法即得(4.4.34)对 $i=1, \cdots, n$ 均成立. 考虑 $i=n$，我们有

$$H_{n+1}Hs_i=s_i \quad (i=1, \cdots, n). \tag{4.4.38}$$

由于 $s_1, \cdots, s_n$ 线性无关，即知 $H_{n+1}H=I$. 所以 $H_{n+1}=H^{-1}$. ∎

Fletcher(1970)发现 Broyden 凸族具有一个很好的性质，如以下定理所述。

定理 4.4.7 设 $f(x)$ 是由 (3.2.1) 给出的，H_1 对称正定，$H_k(k>1)$ 由 Broyden 凸族修正公式给出，令 $\sigma_i^{(k)}=\sigma_i(H^{\frac{1}{2}}H_kH^{\frac{1}{2}})$，即 $H^{\frac{1}{2}}H_kH^{\frac{1}{2}}$ 的第 i 特征值，则有

$$\min\{\sigma_i^{(k)}, 1\}\leqslant\sigma_i^{(k+1)}\leqslant\max\{\sigma_i^{(k)}, 1\}. \tag{4.4.39}$$

§4.5 最小变化性质

拟牛顿法的基本思想是通过逐步修正矩阵 B_k 使其越来越近似 $\nabla^2 f(x)$. 由于每个矩阵 B_k 均在上一次迭代的线搜索方向上满足拟牛顿性质，所以在修正矩阵时，要求 B_{k+1} 与 B_k 之差不要太大. 我们发现不少拟牛顿修正公式具有这样的性质. 即 B_{k+1} 是满足拟牛顿公式(4.2.6)并且在某种范数意义下使得 $B_{k+1}-B_k$ 达到最小.

对任何 s、$y\in\mathbf{R}^n$，我们用 Dennis 和 Schnabel(1983)的记号定义

$$Q(y, s)=\{B\in\mathbf{R}^{n\times n}\,|\,Bs=y\}. \tag{4.5.1}$$

首先，我们有

定理 4.5.1 Broyden 非对称秩为 1 的修正公式(4.3.41)中给出的 B_{k+1} 是问题

$$\min\|B-B_k\|_2 \tag{4.5.2}$$

$$\text{s.t.}\, B\in Q(y_k, s_k) \tag{4.5.3}$$

的解.

证明 设 B_{k+1} 由(4.3.41)给出. 显然，B_{k+1} 满足(4.5.3)，且对任何 $B\in Q(y_k, s_k)$，我们有

$$\|B_{k+1}-B_k\|_2=\left\|\frac{(y_k-B_ks_k)s_k^T}{s_k^Ts_k}\right\|_2=\left\|(B-B_k)\frac{s_ks_k^T}{s_k^Ts_k}\right\|_2$$
$$\leqslant\|B-B_k\|_2. \tag{4.5.4}$$

故知 B_{k+1} 是问题(4.5.2)与(4.5.3)的解. ∎

不难证明，将(4.5.2)中的 $\|\cdot\|_2$ 换成 Frobenins 范数 $\|\cdot\|_F$，定理仍然成立. 定义矩阵范数 $\|\cdot\|_{M,F}$ 为

$$\|A\|_{M,F}=\|MAM\|_F, \tag{4.5.5}$$

其中 M 是 $n\times n$ 对称且非奇异矩阵. Dennis 和 Moré(1977)给出了下面的定理.

定理 4.5.2 设 $B\in\mathbf{R}^{n\times n}$ 对称，s、$y\in\mathbf{R}^n$，$M\in\mathbf{R}^{n\times n}$ 对称非奇异，则问题

$$\min\|A-B\|_{M,F} \tag{4.5.6}$$

$$\text{s.t}\ A\in Q(y,\ s),\ A=A^T \tag{4.5.7}$$

的唯一解是：

$$B_+=B+\frac{(y-Bs)s^TM^{-2}+M^{-2}s(y-Bs)^T}{s^TM^{-2}s}$$
$$+\frac{(y-Bs)^Ts}{(s^TM^{-2}s)^2}M^{-2}ss^TM^{-2}. \tag{4.5.8}$$

证明 对任何 $\bar{B}$ 满足(4.5.7)，我们有

$$M(B_+-B)M=\frac{EZZ^T+ZZ^TE}{Z^TZ}-\frac{Z^TEZ}{(Z^TZ)^2}ZZ^T, \tag{4.5.9}$$

其中 $E=M(\bar{B}-B)M$，$Z=M^{-1}s$. 不难看到

$$\|M(B_+-B)MZ\|_2=\|EZ\|_2, \tag{4.5.10}$$

$$\|M(B_+-B)Mv\|_2=\left\|\frac{ZZ^TEv}{Z^TZ}\right\|_2\leqslant\|Ev\|_2. \tag{4.5.11}$$

其中 v 是任何与 Z 正交的向量. 从(4.5.10)与(4.5.11)即知

$$\|B_+-B\|_{M,F}\leqslant\|E\|_F=\|\bar{B}-B\|_{M,F}. \tag{4.5.12}$$

所以由(4.5.8)定义的 B_+ 是(4.5.6)与(4.5.7)的解. 由于函数 $\varphi(A)=\|A-B\|_{M,F}$ 是严格凸函数，即知 B_+ 是(4.5.6)与(4.5.7)

的唯一解. ▮

定理 4.5.2 的一个直接推论如下:

推论 4.5.3 设 B_k 对称, 由 PSB 公式(4.3.43)给出的 B_{k+1} 是问题

$$\min\|B-B_k\|_F \tag{4.5.13}$$

$$\text{s.t}\, B\in Q(y_k,\ s_k),\ B=B^T \tag{4.5.14}$$

的唯一解.

证明 在定理 4.5.2 中令 $M=I$ 即知推论成立. ▮

另外, 假定 $s_k^Ty_k>0$, 则必存在对称阵 M, 使得 $M^{-2}s_k=y_k$. 于是有

推论 4.5.4 设 B_k 对称, $s_k^Ty_k>0$, 则对任何对称阵 M, 使得 $M^{-2}s_k=y_k$, 问题

$$\min\|B-B_k\|_{M,F} \tag{4.5.15}$$

$$\text{s.t}\, B\in Q(y_k,\ s_k),\ B=B^T \tag{4.5.16}$$

的解是由 DFP 公式(4.3.1)给出的 B_{k+1}.

证明 由定理 4.5.2 知(4.5.15)与(4.5.16)的解为

$$B_+=B_k+\frac{(y_k-B_ks_k)^Ty_k^T+y_k(y_k-B_ks_k)^T}{s_k^Ty_k}$$

$$-\frac{(y_k-B_ks_k)^Ts_k}{(s_k^Ty_k)^2}\,y_ky_k^T. \tag{4.5.17}$$

不难发现(4.5.17)的左端和(4.3.1)的左端完全相等. 所以推论成立. ▮

由于 DFP 公式和 BFGS 公式的对偶性. 我们从推论 4.5.4 即知下面的结果成立.

推论 4.5.5 设 H_k 对称, $s_k^Ty_k>0$, 则对任何对称阵 M, 使得 $M^{-2}s_k=y_k$, 问题

$$\min\|H-H_k\|_{M^{-1},F} \tag{4.5.18}$$

$$\text{s.t.}\ H\in Q(s_k,\ y_k),\ H=H^T \tag{4.5.19}$$

的解是由 BFGS 公式(4.3.10)给出的 H_{k+1}.

设 $y=Gs$ 且 G 是一对称正定阵, 则从(4.5.8)式我们可得到

恒等式

$$M(B_+-G)M=\left(I-\frac{M^{-1}ss^TM^{-1}}{s^TM^{-2}s}\right)$$

$$\times M(B-G)M\left(I-\frac{M^{-1}ss^TM^{-1}}{s^TM^{-2}s}\right).$$

(4.5.20)

为了从(4.5.20)导出一个重要的不等式，我们给出如下引理：

引理 4.5.6　设 A、B 是 $\mathbf{R}^n$ 中两个对称正定阵，s 是 n 维非零实向量．如果

$$B=\left(I-\frac{ss^T}{\|s\|_2^2}\right)A\left(I-\frac{ss^T}{\|s\|_2^2}\right), \tag{4.5.21}$$

则有

$$\|B\|_F^2\leqslant\|A\|_F^2-\frac{\|As\|_2^2}{\|s\|_2^2}. \tag{4.5.22}$$

证明　根据 $\|\cdot\|_F^2$ 的定义和(4.5.21)可得

$$\begin{aligned}\|B\|_F^2&=\mathrm{trace}(BB^T)\\&=\mathrm{trace}(AA^T)-2\frac{s^TA^2s}{s^Ts}+\frac{(s^TAs)^2}{(s^Ts)^2}\\&=\|A\|_F^2-\frac{\|As\|_2^2}{\|s\|_2^2}\\&\quad-\frac{\|As\|_2^2}{\|s\|_2^2}\left(1-\frac{(s^TAs)^2}{\|s\|_2^2\|As\|_2^2}\right)\\&\leqslant\|A\|_F^2-\frac{\|As\|_2^2}{\|s\|_2^2}.\end{aligned} \tag{4.5.23}$$

所以引理成立． ▌

利用引理 4.5.6 和式(4.5.20)可知，如果 B_+ 由 4.5.8 给出，则必有

$$\begin{aligned}\|B_+-G\|_{M,F}^2&\leqslant\|B-G\|_{M,F}^2-\frac{\|M(B-G)s\|_2^2}{\|M^{-1}s\|_2^2}\\&=\|B-G\|_{M,F}^2-\frac{\|M(Bs-y)\|_2^2}{\|M^{-1}s\|_2^2}.\end{aligned}$$

(4.5.24)

上式说明，B_+ 在一定意义下比原矩阵更靠近矩阵 G．不等式

(4.5.24)对于分析拟牛顿方法收敛性也是十分有用的.

另一种要求 B_k 和 B_{k+1} 之间的改变量尽可能小是要求 $B_k^{-1/2}B_{k+1}B_k^{-1/2}$ 尽可能靠近单位矩阵 I. 由于函数

$$\psi(A)=\mathrm{Tr}(A)-\ln(\mathrm{Det}(A)) \tag{4.5.25}$$

对所有正定阵 A 有定义，且在 $A=I$ 时达到最小. 函数(4.5.20)最早由 Byrd 和 Nocedál(1989)提出的. Fletcher(1989) 证明了下列结果:

定理 4.5.7 设 B_k 正定且对称，$s_k^T y_k>0$，则问题

$$\min \psi(B_k^{-\frac{1}{2}}BB_k^{-\frac{1}{2}}) \tag{4.5.26}$$

$$\text{s.t.} \quad B\in Q(y_k,\ s_k),\ B=B^T \tag{4.5.27}$$

的解是由 BFGS 公式(4.3.9)给出的 B_{k+1}.

§4.6 收 敛 性

拟牛顿法的收敛性多年来一直是一个比较热的研究课题. 至今已得到了不少优良结果. 虽然拟牛顿法的收敛性研究已经比较成熟和几乎完善，但仍有一些重要问题还未解决，例如 DFP 方法的总体收敛性问题.

Myers(1968) 首先发现对于凸的二次函数，精确搜索下的 DFP 方法实质上是一个共轭梯度法. 从而知道 DFP 方法具有二次终止性. 现在我们知道，在一定条件下，Huang 族中的所有方法均具有二次终止性(见定理 3.3.4). 利用定理 4.4.7 不难证明对于凸的二次函数，非精确搜索下的 Broyden 凸族中的任何方法均收敛，且

$$\lim_{k\to\infty}\|g_k\|=0.$$

对于一般凸函数，Powell(1971)给出了 DFP 方法的收敛性，如以下定理所述:

定理 4.6.1 设 $f(x)$是一致凸函数且二次连续可微，则由精确线搜索下的 DFP 方法产生的点列$\{x_k\}$必收敛于 $f(x)$ 唯一的极

小点 x^*，且收敛速度是 Q-超线性，即

$$\lim \frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|}=0. \tag{4.6.1}$$

证明 由 DFP 公式(4.3.1)与(4.3.2)可知

$$\mathrm{Tr}(B_{k+1})=\mathrm{Tr}(B_k)-2\frac{y_k^T B_k s_k}{s_k^T y_k}$$

$$+\left(1+\frac{s_k^T B_k s_k}{s_k y_k}\right)\frac{\|y_k\|_2^2}{s_k^T y_k}, \tag{4.6.2}$$

$$\mathrm{Tr}(H_{k+1})=\mathrm{Tr}(H_k)-\frac{\|H_k y_k\|_2^2}{y_k^T H_k y_k}+\frac{\|s_k\|_2^2}{s_k^T y_k}. \tag{4.6.3}$$

因为 $f(x)$ 一致凸且二次连续，必存在常数 $M>0$，使得

$$\frac{\|s_k\|_2^2}{s_k^T y_k}<M,\quad \frac{\|y_k\|_2^2}{s_k^T y_k}<M \tag{4.6.4}$$

对一切 k 均成立．利用 $s_k^T g_{k+1}=0$，可将(4.6.2)表示为

$$\mathrm{Tr}(B_{k+1})=\mathrm{Tr}(B_k)+\frac{\|g_{k+1}\|_2^2}{g_{k+1}^T H_{k+1} g_{k+1}}-\frac{\|g_k\|_2^2}{g_k^T H_k g_k}$$

$$-\frac{\|g_{k+1}\|_2^2}{g_{k+1}^T H_k g_{k+1}}+\frac{\|y_k\|_2^2}{s_k^T y_k}. \tag{4.6.5}$$

由(4.6.3)～(4.6.5)即知存在常数 $M'>0$，使得

$$\mathrm{Tr}(B_{k+1})<\frac{\|g_{k+1}\|_2^2}{g_{k+1}^T H_{k+1} g_{k+1}}-\sum_{j=1}^{k}\frac{\|g_{j+1}\|_2^2}{g_{j+1}^T H_j g_{j+1}}+M'k, \tag{4.6.6}$$

$$\mathrm{Tr}(H_{k+1})\leqslant-\sum_{j=1}^{k}\frac{\|H_j y_j\|_2^2}{y_j^T H_j y_j}+M'k. \tag{4.6.7}$$

不难看出 $g_{j+1}^T H_j g_{j+1}<y_j^T H_j y_j$，于是

$$\sum_{j=1}^{k}\frac{g_{j+1}^T H_j g_{j+1}}{\|y_j\|_2^2}<\sum_{j=1}^{k}\frac{y_j^T H_j y_j}{\|y_j\|_2^2}<\sum_{j=1}^{k}\frac{\|H_j y_j\|_2^2}{y_j^T H_j y_j}<M'k. \tag{4.6.8}$$

利用 Cauchy-Schwarz 不等式即知

$$\sum_{j=1}^{k}\frac{\|g_{j+1}\|_2^2}{g_{j+1}^T H_j g_{j+1}}\geqslant\left(\sum_{j=1}^{k}\frac{\|g_{j+1}\|_2}{\|y_j\|_2}\right)^2\Big/\sum_{j=1}^{k}\frac{g_{j+1}^T H_j g_{j+1}}{\|y_j\|_2^2}$$

$$\geqslant\frac{1}{M'k}\left(\sum_{j=1}^{k}\frac{\|g_{j+1}\|_2}{\|y_j\|_2}\right)^2. \tag{4.6.9}$$

如果 $\{x_k\}$ 不收敛于 $f(x)$ 的唯一的极小点 x^*，则必存在 $\delta>0$，使得

$$\|g_k\|_2 \geqslant \delta \tag{4.6.10}$$

对一切 k 均成立. 另外,由定理 2.5.5 知存在常数 $\eta>0$,使得

$$f(x_k)-f(x_{k+1}) \geqslant \frac{1}{2}\eta\|s_k\|_2^2.$$

从上面不等式即知 $\|s_k\|\to 0$, 故必有 $\|y_k\|\to 0$. 于是从(4.6.9)和(4.6.10)知

$$\sum_{j=1}^{k}\frac{\|g_{j+1}\|_2^2}{g_{j+1}^T H_j g_{j+1}} > M'k \tag{4.6.11}$$

对充分大的 k 都成立. 由(4.6.6)和(4.6.11)可推出

$$\mathrm{Tr}(B_{k+1}) < \frac{\|g_{k+1}\|_2^2}{g_{k+1}^T H_{k+1} g_{k+1}}, \tag{4.6.12}$$

这显然是不可能的. 此矛盾说明了 x_k 必收敛于 x^*. 关于(4.6.1)的证明可参阅 Powell(1971). ▮

Dixon(1972a、b) 把 Huang 族中方法产生相同点列的结果从二次函数推广到一般函数. 下面,我们给出 Broyden 族产生相同点列的简单证明.

定理 4.6.2 设 $f(x)$ 连续可微, $x_1\in\mathbf{R}^n$, $H_1\in\mathbf{R}^{n\times n}$ 给定,则在精确搜索下,Broyden 族中任何方法均产生同一点列.

证明 我们考虑 Broyden 族中的任一方法. 设 $H_k(k>1)$ 由修正公式(4.3.20)产生,且在第 k 次迭代时,$\phi=\phi_k$. 我们把

$$H_{k+1}=H_{k+1}(\phi_1,\ \phi_2,\ \cdots,\ \phi_k) \tag{4.6.13}$$

记成 $\phi_1, \phi_2, \cdots, \phi_k$ 的函数. 首先有

$$H_{k+1}(\phi_1,\ \phi_2,\ \cdots,\ \phi_k)=H_{k+1}(0,\ 0,\ \cdots,\ 0,\ \phi_k), \tag{4.6.14}$$

$$H_{k+1}(\phi_1,\ \phi_2,\ \cdots,\ \phi_k)g_{k+1}=t_k H_{k+1}(0,\ 0,\ \cdots,\ 0)g_{k+1}, \tag{4.6.15}$$

其中

$$t_k=1+\phi_k\frac{g_{k+1}^T H_k(0,\ 0,\ \cdots,\ 0)g_{k+1}}{g_k^T H_k(0,\ 0,\ \cdots,\ 0)g_k}. \tag{4.6.16}$$

对于 $k=1$, 显然(4.6.14)～(4.6.16)成立. 设 $k=1, 2, \cdots, i$ 时,(4.6.14)～(4.6.16)均成立. 于是,有

$$H_{i+2}(\phi_1, \cdots, \phi_{i+1}) = \left(I - \frac{s_{i+1}y_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) H_{i+1}(\phi_1, \cdots, \phi_i)$$
$$\times \left(I - \frac{y_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) + \frac{s_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}} + (\phi_{i+1} - 1) v_{i+1}v_{i+1}^T$$
$$= \left(I - \frac{s_{i+1}y_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) H_{i+1}(0, \cdots, 0, \phi_i)$$
$$\times \left(I - \frac{y_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) + \frac{s_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}} + (\phi_{i+1} - 1) v_{i+1}v_{i+1}^T$$
$$= \left(I - \frac{s_{i+1}y_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) H_{i+1}(0, \cdots, 0)$$
$$\times \left(I - \frac{y_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}}\right) + \frac{s_{i+1}s_{i+1}^T}{s_{i+1}^T y_{i+1}} + (\phi_{i+1} - 1) v_{i+1}v_{i+1}^T$$
$$= H_{i+2}(0, \cdots, 0, \phi_{i+1}). \tag{4.6.17}$$

故知

$$H_{i+2}(\phi_1, \cdots, \phi_{i+1}) g_{i+2} = H_{i+2}(0, \cdots, 0, 0)$$
$$\times g_{i+2} + \phi_{i+1} v_{i+1} v_{i+1}^T g_{i+2}$$
$$= \left(1 + \phi_{i+1} \frac{g_{i+2}^T H_{i+1}(0, \cdots, 0) g_{i+2}}{g_{i+1}^T H_{i+1}(0, \cdots, 0) g_{i+1}}\right)$$
$$\times H_{i+2}(0, \cdots, 0) g_{i+2}. \tag{4.6.18}$$

由(4.6.17)和(4.6.18)可知(4.6.14)～(4.6.16)对 $k=i+1$ 也成立．由归纳法知(4.6.14)～(4.6.16)对一切 k 都成立．所以，只要我们假定对一切 k, $t_k \neq 0$ 则从(4.6.15)可知 Broyden 族中任何方法均产生相同方向，从而必产生相同的点列．所以定理成立．▌

Burmeister(1973) 以及 Schuller 和 Stoer(1974) 独立地证明了 DFP 方法以及整个 Broyden 族在精确线搜索下的 n 步二次收敛性，即存在常数 c，使得

$$\|x_{k+n} - x^*\|_2 \leq c\|x_k - x^*\|_2^2 \tag{4.6.19}$$

对一切 k 都成立．由于 Broyden 族中任何方法在精确线搜索时产生相同的点列，所以考虑 Broyden 族中的方法在精确线搜索时的收敛速度只需讨论其中任何一个就够了．Ritter(1980) 将(4.6.19)改进到

$$\|x_{k+n} - x^*\|_2 = o(\|x_k - x^*\|_2^2). \tag{4.6.20}$$

如果我们定义矩阵 U_k 如下：

$$U_k=[d_{k+1}/\|d_{k+1}\|_2, \cdots, d_{k+n}/\|d_{k+n}\|_2] \qquad (4.6.21)$$

Schuller(1974)在假定

$$\limsup_{k\to\infty}\|U_k^{-1}\|<+\infty \qquad (4.6.22)$$

下证明了

$$\|x_{k+n}-x^*\|=O(\|x_{k+n-1}-x^*\|\|x_{k-1}-x^*\|). \qquad (4.6.23)$$

Ritter(1980)在(4.6.22)的假定下将结果(4.6.23)改进到

$$\|x_{k+n}-x^*\|=O(\|x_{k+n-1}-x^*\|\|x_k-x^*\|). \qquad (4.6.24)$$

不难构造出例子来使得所有的 U_k 都是奇异阵，因而(4.6.22)是一个极强的假定. 在不假定(4.6.22)的情况下，Powell(1984b)证明了

$$\|x_{k+n}-x^*\|=O(\|x_{k+1}-x^*\|\|x_k-x^*\|). \qquad (4.6.25)$$

当 $n=2$ 时，估计式(4.6.25)和(4.6.24)是一样的，而且此结果是最优的. 当 $n>2$ 时，(4.6.25)比(4.6.24)要弱. 尽管大量的数值计算表明了(4.6.24)在没有(4.6.22)的假定也是成立的，但从理论上证明它是否对，仍是一个没有解决的难题.

另外，考虑变尺度算法的 Q-阶收敛速度与牛顿方法的 Q-阶收敛速度之间的关系是很有意义的. 由 Ortega 和 Rheinboldt 1970)可知，不进行线搜索的牛顿方法(即在算法 4.1.1 中取 $\alpha_k=1$) 的 Q-收敛阶为 $1+\varepsilon$，其中 $\varepsilon>0$ 是目标函数 $f(x)$ 的海色阵的 Hölder 连续因子. Powell(1984b) 给出一个三次连续可微的目标函数只有小于2的 Q-阶收敛，从而说明拟牛顿算法的 Q-收敛阶可能小于 $1+\varepsilon$. Yuan(1984a) 证明了对于所有的二次可微函数，变尺度算法的最小 Q-收敛阶为 1. 作者认为以下的估计式

$$\frac{\|x_k-x^*\|_2}{\|x_{k-1}-x^*\|_2}=O(\|x_{k-n}-x^*\|_2^{\frac{\varepsilon}{2}}) \qquad (4.6.26)$$

可能是在精确线搜索下 Broyden 族方法的收敛速度. 但至今我们还没能从理论上证明(4.6.26)正确与否.

另一种局部收敛性分析是讨论没有线搜索的拟牛顿方法的收

敛速度. 这里没有线搜索是指在每次迭代都令 $\alpha_k=1$. 研究这种特殊情形是有实际背景的. 第一, 几乎所有的线搜索方法都是先试步长为 1; 第二, 当点列 x_k 收敛于解时, 不少拟牛顿法(如 DFP 方法, BFGS 方法)取步长为 1 得到的点将满足非精确线搜索条件 (2.5.8) 与 (2.5.9). 在这方面, 不少重要的结果由 Dennis 和 Moré(1977)给出, 我们列出其中之一. 如以下定理所述:

定理 4.6.3 假定 $f(x)$ 二次连续可微, 且 $\nabla^2 f(x)$ 满足 Lipschitz 条件. 设 x^* 是 $f(x)$ 的极小点, 且 $\nabla^2 f(x^*)$ 正定, 则 DFP 方法和 BFGS 方法在取 $\alpha_k\equiv 1$ 时是局部 Q-超线性收敛于 x^* 的.

定理 4.6.3 证明的基本思想是先证明点列线性收敛, 从而有

$$\sum_{k=1}^{\infty}\|x_k-x^*\|_2<+\infty. \tag{4.6.27}$$

利用(4.6.27)即可证明 B_k 和 H_k 一致有界, 且

$$\lim\frac{\|(B_k-\nabla^2 f(x^*))s_k\|_2}{\|s_k\|_2}=0. \tag{4.6.28}$$

由于 B_k 和 H_k 一致有界, 即知

$$m\leqslant\|g_k\|_2/\|d_k\|_2\leqslant M. \tag{4.6.29}$$

由于 $(B_k-\nabla^2 f(x^*))d_k=-\nabla f(x_k+d_k)+o(\|d_k\|)$, 从(4.6.28)即知

$$\lim_{k\to\infty}\frac{\|\nabla f(x_k+d_k)\|_2}{\|d_k\|_2}=0. \tag{4.6.30}$$

利用(4.6.29)和(4.6.30)可证明

$$\lim\frac{\|x_k+d_k-x^*\|_2}{\|x_k-x^*\|_2}=0. \tag{4.6.31}$$

故知步长为 1 可使算法 Q-超线性收敛.

不难发现, 将 $\alpha_k\equiv 1$ 换成 $\alpha_k\to 1$, 定理 4.6.3 仍然成立.

关于非精确线搜索下拟牛顿法的全局收敛性, Powell(1976 a) 开创性地证明了 BFGS 方法的收敛性, 如下面定理所述:

定理 4.6.4 设 $f(x)$ 是凸函数, 集合 $\{x;\ f(x)\leqslant f(x_1)\}$ 有界, 且 $f(x)$ 在该集合上二次连续可微, B_1 是任给的正定矩阵, 则 BFGS 方法在非精确线搜索下产生的点列 x_k 必有限终止, 或者

$$\lim_{k\to\infty}\|g_k\|_2=0. \tag{4.6.32}$$

证明 根据假设，必存在常数 $M>0$，使得

$$\frac{\|y_k\|_2^2}{s_k^T y_k}\leqslant M \tag{4.6.33}$$

对一切 k 都成立．于是，由修正公式(4.3.9)即知

$$\begin{aligned}\operatorname{Tr}(B_{k+1})&=\operatorname{Tr}(B_1)-\sum_{i=1}^{k}\frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k}+\sum_{i=1}^{k}\frac{\|y_k\|_2^2}{s_k^T y_k}\\&\leqslant-\sum_{i=1}^{k}\frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k}+\overline{M}k.\end{aligned} \tag{4.6.34}$$

其中 $\overline{M}=M+\operatorname{Tr}(B_1)$．从(4.6.34)可推出

$$\operatorname{Det}(B_{k+1})\leqslant[\operatorname{Tr}(B_{k+1})/n]^n\leqslant[\overline{M}k/n]^n \tag{4.6.35}$$

以及

$$\sum_{i=1}^{k}\frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k}\leqslant\overline{M}k. \tag{4.6.36}$$

由(4.6.36)可知

$$\prod_{i=1}^{k}\frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k}\leqslant\overline{M}^k. \tag{4.6.37}$$

从(4.3.11)可得到关系式

$$\prod_{i=1}^{k}\frac{s_k^T y_k}{s_k^T B_k s_k}=\frac{\operatorname{Det}(B_{k+1})}{\operatorname{Det}(B_1)}. \tag{4.6.38}$$

利用(4.6.35)、(4.6.37)和(4.6.38)，我们有

$$\prod_{i=1}^{k}\frac{\|B_k s_k\|_2^2 s_k^T y_k}{(s_k^T B_k s_k)^2}\leqslant\frac{[\overline{M}k/n]^n\overline{M}^k}{\operatorname{Det}(B_1)}\leqslant(\hat{M})^k. \tag{4.6.39}$$

其中 $\hat{M}>0$ 是某一常数．由非精确线搜索条件(2.5.9)知

$$s_k^T y_k\geqslant-(1-b_2)s_k^T g_k. \tag{4.6.40}$$

从而利用(4.6.40)和(4.6.39)可证

$$\prod_{i=1}^{k}\frac{\|g_k\|_2^2}{-s_k^T g_k}\leqslant\left(\frac{\hat{M}}{1-b_2}\right)^k. \tag{4.6.41}$$

由非精确线搜索条件(2.5.8)知

$$\sum_{i=1}^{\infty}-s_k^T g_k<+\infty. \tag{4.6.42}$$

于是

$$\lim_{k\to\infty} -s_k^T g_k = 0. \tag{4.6.43}$$

由(4.6.41)和(4.6.43)即知

$$\liminf_{k\to\infty} \|g_k\|_2^2 = 0. \tag{4.6.44}$$

因为 $f(x)$是凸函数, 显然, (4.6.44)等价于(4.6.32). ▌

定理 4.6.4 由 Byrd、Nocedal 和 Yuan(1987)推广到除 DFP 以外 Broyden 凸族中的所有方法.

定理 4.6.5 在定理 4.6.4 的假定下, 设 B_{k+1} 由 Broyden 凸族(4.3.14)给出, 即 $B_{k+1}=B_{k+1}(\theta_k)$, 若存在 $\delta>0$, 使得 $1-\theta_k\geqslant\delta$, 则在非精确线搜索下的算法 4.2.1 必有限终止, 或者

$$\lim \|g_k\|_2 = 0. \tag{4.6.45}$$

证明 假定定理不真, 则存在常数 $\delta'>0$, 使得

$$\|g_k\|_2 \geqslant \delta' \tag{4.6.46}$$

对一切 k 都成立. 由于

$$\mathrm{Tr}(B_{k+1}) = \mathrm{Tr}(B_k) - (1-\theta_k)\frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k} + \left(1+\theta_k\frac{s_k^T B_k s_k}{s_k^T y_k}\right)\frac{\|y_k\|_2^2}{s_k^T y_k} - 2\theta_k\frac{y_k^T B_k s_k}{s_k^T y_k} \tag{4.6.47}$$

以及

$$\frac{\|y_k\|^2}{s_k^T y_k}\frac{s_k^T B_k s_k}{s_k^T y_k} \div \frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k} \leqslant M\frac{(s_k^T B_k s_k)^2}{s_k^T y_k\|B_k s_k\|_2^2} = M\frac{(s_k^T g_k)^2}{s_k^T y_k\|g_k\|_2^2} \leqslant M(1-b_2)^{-1}(-s_k g_k)/\|g_k\|_2^2. \tag{4.6.48}$$

$$\frac{|y_k^T B_k s_k|}{s_k^T y_k} \div \frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k} \leqslant \frac{\|y_k\| s_k^T B_k s_k}{y_k^T s_k\|B_k s_k\|_2} \leqslant \sqrt{M}\frac{s_k^T B_k s_k}{\sqrt{s_k^T y_k}\,\|B_k s_k\|_2} \leqslant \sqrt{M(1-b_2)^{-1}}\sqrt{-s_k^T g_k}/\|g_k\|_2. \tag{4.6.49}$$

由(4.6.43)、(4.6.46)～(4.6.49)即知

$$\operatorname{Tr}(B_{k+1}) \leqslant \operatorname{Tr}(B_k) - \frac{\delta}{2} \frac{\|B_k s_k\|_2^2}{s_k^T B_k s_k} + \frac{\|y_k\|_2^2}{s_k^T y_k} \tag{4.6.50}$$

对所有充分大的 k 均成立. 由于 $\theta_k \geqslant 0$, 故知

$$\prod_{i=1}^{k} \frac{s_k^T y_k}{s_k^T B_k s_k} \leqslant \frac{\operatorname{Det}(B_{k+1})}{\operatorname{Det}(B_1)}. \tag{4.6.51}$$

现在利用(4.6.50)和(4.6.51), 与定理 4.6.4 的证明类似, 我们可证

$$\lim_{k\to\infty} \|g_k\|_2 = 0. \tag{4.6.52}$$

这与假定(4.6.46)相矛盾. 此矛盾说明了定理成立. ▌

目前, 人们仍不知道在非精确线搜索下, DFP 方法是否一定收敛. 这是有待解决的一个重要问题.

另外, 对于非凸函数的情形, 至今仍没有任何总体收敛性结果. 分析非凸目标函数的一个主要难点是(4.6.33) 一般不成立, 所以即使对于 BFGS 方法也很难证明 (4.6.34) 成立. 在假定点列 x_k 收敛于 x^* 的前提下, Powell(1972)证明了如果 $n=2$, 且线搜索为精确搜索, 则 DFP 方法必收敛于 $f(x)$ 的稳定点. 后来濮定国和俞文鮆(1988)将 Powell 的结果推广到 $n\geqslant 2$.

分析算法的全局收敛性, 除了 Powell 提出的研究矩阵 B_k 和 H_k 的追迹以及行列式的值的方法以外, 最近 Byrd 和 Nocedal (1989)提出了用估计 $\psi(B_k)$ 来研究算法的收敛性. 这里 $\psi(\cdot)$ 函数由(4.5.20)定义. 而且利用 $\psi(\cdot)$ 函数同时可进行超线性收敛分析.

对拟牛顿法的局部超线性收敛的分析的一类重要方法是估计 $\|B_{k+1}-B_k\|$, 最早可见 Broyden、Dennis 和 Moré(1973). 实质上, 这类方法的思想是基于修正矩阵的最小变化性质. 关于这方面的较新结果可见 Martínez(1990). 超线性收敛分析一般归结到证明 Dennis 和 Moré(1974) 给出的超线性收敛特征, 即(4.6.28)成立和

$$\alpha_k \to 1. \tag{4.6.53}$$

Byrd Liu 和 Nocedal(1990)将(4.6.28)换成等价条件

$$\lim_{k\to\infty}\cos\langle\tilde{d}_k, -\tilde{g}_k\rangle=\lim_{k\to\infty}\frac{s_k^TB_ks_k}{s_k^T\nabla^2f(x^*)s_k}=1. \quad (4.6.54)$$

其中 $\tilde{d}_k=[\nabla^2f(x^*)]^{\frac{1}{2}}d_k$, $\tilde{g}_k=[\nabla^2f(x^*)]^{-\frac{1}{2}}g_k$. 事实上，超线性收敛等价于算法十分接近精确线搜索的牛顿法.

定理 4.6.6 设迭代公式

$$x_{k+1}=x_k-\alpha_kB_k^{-1}g_k \quad (4.6.55)$$

产生的点列收敛于 x^*, $\nabla^2f(x^*)$正定, $\nabla f(x^*)=0$, 则

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x^*\|_2}{\|x_k-x^*\|_2}=0 \quad (4.6.56)$$

当且仅当

$$\lim_{k\to\infty}\cos^2\langle B_k^{-1}g_k, -[\nabla^2f(x^*)]^{-1}g_k\rangle=1 \quad (4.6.57)$$

和

$$\lim_{k\to\infty}\frac{s_k^TB_ks_k}{\alpha_ks_k^Ty_k}=1. \quad (4.6.58)$$

证明 假定(4.6.56)成立，我们有

$$\lim_{k\to\infty}\cos^2\langle B_k^{-1}g_k, x_k-x^*\rangle=1, \quad (4.6.59)$$

显然有关系式

$$\lim_{k\to\infty}\cos^2\langle x_k-x^*, -[\nabla^2f(x^*)]^{-1}g_k\rangle=1. \quad (4.6.60)$$

由 (4.6.59) 和 (4.6.60) 即知 (4.6.57) 成立. 从 (4.6.56) 和 $\nabla^2f(x^*)$正定还可知

$$\lim_{k\to\infty}\frac{\|g_k+y_k\|_2}{\|g_k\|_2}=0, \quad (4.6.61)$$

于是

$$\lim_{k\to\infty}\frac{s_k^Tg_k+s_k^Ty_k}{\|s_k\|\|g_k\|}=0. \quad (4.6.62)$$

从(4.6.62)即知

$$\lim_{k\to\infty}\frac{-s_k^Tg_k}{s_k^Ty_k}=1. \quad (4.6.63)$$

故知(4.6.58)成立.

现假定(4.6.57)和(4.6.58)成立. 由(4.6.57)和(4.6.60)知

(4.6.59)成立．从(4.6.58)可知(4.6.63)成立．于是

$$\lim_{k\to+\infty}\frac{s_k^T g_k+s_k\nabla^2 f(x^*)s_k}{s_k^T y_k}=0. \qquad (4.6.64)$$

所以

$$\lim_{k\to\infty}\frac{s_k^T\nabla^2 f(x^*)[s_k+(\nabla^2 f(x^*))^{-1}g_k]}{s_k^T\nabla^2 f(x^*)s_k}=0. \qquad (4.6.65)$$

由(4.6.65)和(4.6.57)即知

$$\lim_{k\to\infty}\frac{\|s_k+(\nabla^2 f(x^*))^{-1}g_k\|}{\|s_k\|}=0. \qquad (4.6.66)$$

上式显然等价于(4.6.56)．所以定理成立． ▌

§4.7 有限内存 BFGS 方法

对于中小规模的无约束优化问题，拟牛顿法(如 BFGS 方法)是十分有效的．但对于大规模问题，即 n 相当大时，算法所需内存相当重要，并且在每次迭代中线代数计算量也影响算法的效率．

有限内存 (limited Memory) 拟牛顿法可看成是共轭梯度法的推广，在 3.4 节中曾提到共轭梯度法和拟牛顿法之间的联系．有限内存拟牛顿法最早由 Perry(1977) 和 Shanno(1978) 提出的，此后有不少人对其进行研究，如 Gill 和 Murray(1979)、Buckley (1978)、Buckley 和 LeNir(1983) 以及 Nocedal(1980)．

有限内存 BFGS 方法的基本出发点是减少内存．由于 BFGS 修正公式(4.3.10)可写成

$$H_{k+1}=\left(I-\frac{s_k y_k^T}{s_k^T y_k}\right)H_k\left(I-\frac{y_k s_k^T}{s_k^T y_k}\right)+\frac{s_k s_k^T}{s_k^T y_k}. \qquad (4.7.1)$$

记 $\rho_k=1/s_k^T y_k$ 以及

$$V_k=(I-\rho_k y_k s_k^T), \qquad (4.7.2)$$

则

$$H_{k+1}=(V_k^T\cdots V_{k-i}^T)H_{k-i}(V_{k-i}\cdots V_k)$$

$$+\sum_{j=0}^{i}\rho_{k-i+j}\left(\prod_{l=0}^{i-j-1}V_{k-l}^{T}\right)s_{k-i+j}s_{k-i+j}^{T}\left(\prod_{l=0}^{i-j-1}V_{k-l}^{T}\right)^{T}. \tag{4.7.3}$$

所以 $m+1$ 步的有限 BFGS 方法正是利用

$$H_{k+1}=V_k^T\cdots V_{k-m}^T H_k^{(0)} V_{k-m}\cdots V_k$$
$$+\sum_{j=0}^{m}\rho_{k-m+j}\prod_{l=0}^{m-j-1}V_{k-l}^{T}s_{k-m+j}s_{k-m+j}^{T}\left(\prod_{l=0}^{m-j-1}V_{k-l}^{T}\right)^{T}. \tag{4.7.4}$$

其中 $H_k^{(0)}$ 是一个预先给定的简单正定阵或者由某种方式自动产生. 我们发现，在(4.7.4)中，由于假定 $H_k^{(0)}$ 已知，我们只需存储 s_i、$y_i(i=k-m,\ \cdots,\ k)$就够了. $H_k^{(0)}$ 的一种选取方式为

$$H_k^{(0)}=\frac{s_k^T y_k}{\|y_k\|_2^2}I. \tag{4.7.5}$$

关于其它选取 $H_k^{(0)}$ 的方法，请参阅 Liu 和 Nocedal(1989). 于是我们得到有限内存 BFGS 算法如下:

算法 4.7.1

步 1　给出 $x_1\in\mathbf{R}^n$, $H_1\in\mathbf{R}^{n\times n}$ 对称正定;

取非负整数 m;

取 $0<b_1\leqslant b_2<1$; $\varepsilon\geqslant 0$;

$k:=1$

步 2　如果 $\|g_k\|\leqslant\varepsilon$, 则停;

计算 $d_k=-H_k g_k$.

步 3　利用非精确线搜索(2.5.8)与(2.5.9)

求 $\alpha_k>0$;

$$x_{k+1}=x_k+\alpha_k d_k. \tag{4.7.6}$$

步 4　$\hat{m}=\min\{k,\ m\}$,

如果 $k=1$, 则 $H_1^{(0)}=H_1$, 否则, $H_k^{(0)}$ 由(4.7.5)给出;

$$H_{k+1}=(V_k^T\cdots V_{k-\hat{m}}^T)H_k^{(0)}V_{k-\hat{m}}\cdots V_k$$
$$+\sum_{j=0}^{\hat{m}}\rho_{k-\hat{m}+j}\left(\prod_{l=0}^{\hat{m}-j-1}V_{k-l}^{T}\right)s_{k-\hat{m}+j}s_{k-\hat{m}+j}^{T}\left(\prod_{l=0}^{\hat{m}-j-1}V_{k-l}\right)^{T}; \tag{4.7.7}$$

步 5 $k:=k+1$; 转步 2.

对于一致凸函数，不难发现 B_k 和 H_k 一致有界，从而算法必收敛。

定理 4.7.2 设函数二次连续可微且一致凸，由算法 4.7.1 产生的点列$\{x_k\}$收敛于$f(x)$唯一的极小点，且收敛速度至少是 R-线性的。

证明 由于$f(x)$一致凸，故点列$\{x_k\}$必有界且存在常数 M，使得

$$\frac{\|y_k\|_2^2}{s_k^T y_k} \leqslant M, \tag{4.7.8}$$

$$\frac{\|s_k\|_2^2}{s_k^T y_k} \leqslant M. \tag{4.7.9}$$

由(4.7.8)和(4.7.9)即知

$$\|V_k\|_2 \leqslant 1+M. \tag{4.7.10}$$

不失一般性，可假定$\|H_1\|_2 \leqslant M$. 于是利用(4.7.7)～(4.7.10)，即知

$$\|H_{k+1}\|_2 \leqslant (m+1)(1+M)^{2(m+1)} M \tag{4.7.11}$$

对一切 k 均成立. 另一方面，从(4.7.7)可知

$$B_k^{(i+1)} = B_k^{(i)} - \frac{B_k^{(i)} s_{k-\hat{m}+i} s_{k-\hat{m}+i}^T B_k^{(i)}}{s_{k-\hat{m}+i}^T B_k^{(i)} s_{k-\hat{m}+i}} + \frac{y_{k-\hat{m}+i} y_{k-\hat{m}+i}^T}{s_{k-\hat{m}+i}^T y_{k-\hat{m}+i}}. \tag{4.7.12}$$

$B_k^{(0)} = [H_k^{(0)}]^{-1}$，则 $B_k^{(\hat{m}+1)} = B_{k+1} = H_{k+1}^{-1}$. 利用(4.7.5)、(4.7.8)和(4.7.12)可证

$$\mathrm{Tr}(B_{k+1}) \leqslant mM. \tag{4.7.13}$$

由(4.7.11)和(4.7.13)知存在 $\delta>0$，使得

$$\cos\langle d_k, -g_k\rangle \geqslant \delta > 0 \tag{4.7.14}$$

对一切 k 都成立. 从(4.7.14)就容易证明存在 $\delta' \in (0, 1)$，使得(见 Powell, 1976a)：

$$f(x_{k+1}) - f(x^*) \leqslant \delta'(f(x_k) - f(x^*)). \tag{4.7.15}$$

由一致凸性，存在 $\bar{\delta}>0$，使得

$$\|x_k-x^*\|\leqslant\bar{\delta}[f(x_k)-f(x^*)]^{1/2}$$
$$\leqslant\bar{\delta}(\delta')^{-1}[f(x_1)-f(x^*)]^{1/2}(\sqrt{\delta'})^k. \tag{4.7.16}$$

(4.7.16)表明 x_k R-线性收敛于 x^*. ▮

当 $m=0$, 算法 4.7.1 实质上就是一个共轭梯度法. 由于一般总有 $m\ll n$, 所以看来算法 4.7.1 不可能有超线性收敛结果.

对于 $m=0$, 我们可将定理 4.7.2 的结果推广到一般凸函数.

定理 4.7.3 设函数 $f(x)$ 是连续可微的凸函数, 且 $\nabla f(x)$ 满足 Lipschitz 条件

$$\|\nabla f(y)-\nabla f(z)\|\leqslant M\|y-z\|. \tag{4.7.17}$$

则由算法 4.7.1($m=0$)产生的点列 x_k 必满足

$$\lim_{k\to\infty}f(x_k)=-\infty \tag{4.7.18}$$

或者

$$\lim_{k\to\infty}\|g_k\|=0. \tag{4.7.19}$$

证明 用反证法: 假定定理不真, 则必有

$$\sum_{k=1}^{\infty}[f(x_k)-f(x_{k+1})]<+\infty, \tag{4.7.20}$$

$$\lim_{k\to\infty}\sup\|g_k\|_2>0. \tag{4.7.21}$$

由于 $f(x)$ 是凸函数, 故由(4.7.21)可知, 存在常数 $\delta>0$, 使得

$$\|g_k\|_2\geqslant\delta \tag{4.7.22}$$

对一切 k 均成立. 因为 $m=0$, 我们有

$$B_{k+1}=\frac{\|y_k\|_2^2}{s_k^Ty_k}\left(I-\frac{s_ks_k^T}{s_k^Ts_k}\right)+\frac{y_ky_k^T}{s_k^Ty_k}, \tag{4.7.23}$$

$$H_{k+1}=\frac{s_k^Ty_k}{\|y_k\|_2^2}I-\frac{s_ky_k^T+y_ks_k^T}{\|y_k\|_2^2}+2\frac{s_ks_k^T}{s_k^Ty_k}. \tag{4.7.24}$$

所以

$$\mathrm{Tr}(B_{k+1})=n\frac{\|y_k\|_2^2}{s_k^Ty_k}, \tag{4.7.25}$$

$$\mathrm{Tr}(H_{k+1})=(n-2)\frac{s_k^Ty_k}{\|y_k\|_2^2}+2\frac{\|s_k\|_2^2}{s_k^Ty_k}\leqslant n\frac{\|s_k\|_2^2}{s_k^Ty_k}. \tag{4.7.26}$$

由非精确线搜索条件(2.5.9)知

$$s_k^T y_k > (1-b_2)\alpha_k g_k^T H_k g_k. \tag{4.7.27}$$

由假定(4.7.17)知

$$\frac{\|y_k\|_2^2}{s_k^T y_k} \leqslant M. \tag{4.7.28}$$

利用(4.7.20)、(2.5.8)、(4.7.22)、(4.7.25)和(4.7.28)可得:

$$\begin{aligned} +\infty > \sum_{k=1}^{\infty} [f(x_k) - f(x_{k+1})] \geqslant b_1 \sum_{k=1}^{\infty} -s_k^T g_k \\ = b_1 \sum_{k=1}^{\infty} \alpha_k g_k^T H_k g_k \geqslant b_1 \delta^2 \sum_{k=1}^{\infty} \frac{\alpha_k}{\mathrm{Tr}(B_k)} \\ \geqslant b_1 \delta^2 (nM)^{-1} \sum_{k=1}^{\infty} \alpha_k. \end{aligned} \tag{4.7.29}$$

于是, $\alpha_k \to 0$. 从(4.7.26)和(4.7.28)可知

$$\mathrm{Tr}(H_{k+1}) \leqslant n \frac{\alpha_k g_k^T H_k^2 g_k}{(1-b_2) g_k^T H_k g_k} \leqslant \frac{n\alpha_k}{(1-b_2)} \mathrm{Tr}(H_k). \tag{4.7.30}$$

从而必有 $\mathrm{Tr}(H_k) \to 0$. 这与(4.7.25) 矛盾. 此矛盾说明了定理为真. ∎

定理 4.7.3 对于 $m>0$ 是否成立, 仍是一个尚待解决的问题. 另外, 取不同值 m 时, 对算法效率在理论上的比较, 尚无人进行分析过.

在实际计算中, 显然 m 的大小取决于问题的维数、机器容许的内存. 一般说来, m 的取值在 3 至 8 之间(见 Liu 和 Nocedal, 1989).

§4.8 修正公式的几种计算形式

在拟牛顿算法中, 每次迭代的线搜索方向 $d_k = -B_k^{-1} g_k$. 如果算法利用修正公式产生 B_k, 则每次迭代需解线性方程组

$$B_k d = -g_k. \tag{4.8.1}$$

解线性方程组需要 $O(n^3)$ 计算量. 所以人们常利用逆矩阵的修正公式产生 H_k, 于是只需计算

$$d_k := H_k g_k. \tag{4.8.2}$$

计算(4.8.2)的右端只需 $O(n^2)$ 计算量.但美中不足的是:修正 H_k 的数值稳定性往往比修正 B_k 要差得多.一个简单的解释是,修正 B_k 需要计算 $y_k y_k^T / s_k^T y_k$;而修正 H_k 需要计算 $s_k s_k^T / s_k^T y_k$.假定 s_k、y_k 均有计算误差 ε,则计算 $y_k y_k^T / s_k^T y_k$ 和 $s_k s_k^T / s_k^T y_k$ 的误差分别为 $O([\|y_k\|_2^2 / s_k^T y_k] \cdot \varepsilon)$ 和 $O(\varepsilon \|s_k\|_2^2 / s_k^T y_k)$.不难看出,对于二次连续可微函数,且 $\nabla f(x)$ 满足 Lipschitz 条件,则存在常数 M,使得 $\|y_k\|_2^2 < M\|s_k\|_2^2$.而反过来,有可能 $\|s_k\|/\|y_k\| \to \infty$.

为了使算法既基于修正 B_k,又避免解方程组(4.8.1),Gill 和 Murray(1972)提出把 B_k 分解成 LDL^T 的形式,其中 L 是单位下三角矩阵,D 是对角阵.然后每次迭代修正 L 和 D.由于秩 2 修正公式均可写成

$$B_{k+1} = B_k + \tau_k u_k u_k^T + \xi_k z_k z_k^T, \tag{4.8.3}$$

其中 τ_k、$\xi_k \in \mathbf{R}$,u_k、$z_k \in \mathbf{R}^n$.所以只需要考虑

$$LDL^T + \tau Z Z^T \tag{4.8.4}$$

的 LDL 分解即可.Gill 和 Murray(1972)是将(4.8.4)中矩阵写成

$$L(D + \tau \bar{Z} \bar{Z}^T) L. \tag{4.8.5}$$

其中 $\bar{Z}$ 由解下三角方程组 $L\bar{Z} = Z$ 得到.再考虑一个对角阵加上一个秩为 1 的阵的 LDL^T 分解.利用这方法求得(4.8.4)的 LDL^T 分解需要的计算量约为 $\frac{3}{2}n^2$.所以,已知 $B_k = L_k D_k L_k^T$,求出(4.8.3)中 B_{k+1} 的 LDL^T 分解所需计算量约为 $3n^2$.这种利用修正 B 的 LDL^T 分解的方法在每次迭代只需求解

$$L_k D_k L_k^T d = -g_k \tag{4.8.6}$$

来确定搜索方向.求解(4.8.6)只需解两个三角方程组,故只需 $O(n^2)$ 计算量.修正 B 的 LDL^T 分解的另一个优点是只要 D_k 的对角线元素全为正,则 B_k 必正定.但它的缺点是程序复杂.

最近,Powell(1987)提出基于 H_k 的对称分解的一种修正公式在每次迭代时有

$$H_k = Z_k Z_k^T, \tag{4.8.7}$$

其中 Z_k 是一非奇异阵. 已知 Z_k, 则显然线搜索方向 $d_k=-H_k g_k$ 在 $O(n^2)$ 次计算量内可求得. 然后, Powell 用 BFGS 公式求 H_{k+1} 的分解形式 $Z_{k+1}Z_{k+1}^T$. 不难发现, BFGS 公式(4.3.10)可写成

$$H_{k+1}=(I-s_kP_k^T)H_k(I-P_ks_k^T). \tag{4.8.8}$$

其中

$$P_k=\frac{B_ks_k}{\sqrt{s_k^Ty_ks_k^TB_ks_k}}-\frac{y_k}{s_k^Ty_k}. \tag{4.8.9}$$

所以我们可取

$$Z_{k+1}=(I-s_kP_k^T)Z_kQ_k, \tag{4.8.10}$$

其中 Q_k 可以是任何正交矩阵. 从(4.8.7)式可知

$$Z_k^TB_kZ_k=I. \tag{4.8.11}$$

所以 Z_k 的列向量 $z_k^{(i)}$ 是 B_k-共轭的. 于是, Powell(1987)选取正交阵 Q_k, 使得 Z_kQ_k 的第一列是和 s_k 在一个方向上. 取 Q_k 为下 Hessenberg 矩阵, 则可证明对于凸的二次函数在精确线搜索下有

$$z_{k+1}^{(i)}Hz_{k+1}^{(j)}=0(1\leqslant i\leqslant k<j\leqslant n), \tag{4.8.12}$$

其中 H 是函数 $f(x)$ 的海色阵. 由于共轭性质(4.8.12), 则该算法显然具有二次终止性.

Powell(1987)还提出对 $z_{k+1}^{(i)}$ 进行加权. 设 $Z_{k+1}=(z_{k+1}^{(1)},\cdots,z_{k+1}^{(n)})$ 已求得, 我们对它的第 2 到 n 列进行加权:

$$z_{k+1}^{(i)}:=\max\left\{1,\ \frac{\sigma_k}{\|z_{k+1}^{(i)}\|_2}\right\}z_{k+1}^{(i)}, \tag{4.8.13}$$

其中 $\sigma_1=\|z_2^{(1)}\|_2$, 且

$$\sigma_k=\min\{\sigma_{k-1},\ \|z_{k+1}^{(1)}\|_2\}. \tag{4.8.14}$$

加权以后的 Z_{k+1} 显然不再满足 BFGS 修正公式, 但是 $z_{k+1}^{(i)}$ 的方向仍然保持, 故这个加权的 BFGS 方法仍具有二次终止性. 但对于一般非线性函数, 这个方法的收敛性尚无人进行分析.

基于 Powell(1987)的方法, Lalee 和 Nocedal(1991)提出一列自调节 (automatic column scaling)BFGS 方法, 并在非精确搜索的假定下证明了该方法对于一致凸函数是收敛的, 而且局部收敛速度是 Q-超线性的.

第 5 章

直接方法

非线性规划的直接方法是指：不利用函数任何导数的方法．本章讨论无约束优化问题

$$\min_{x\in\mathbf{R}^n} f(x) \tag{5.0.1}$$

的直接方法．我们介绍几个早期的直接方法，如 Hooke 和 Jeeves 方法、Rosenbrock 方法和单纯形法．然后讨论目前广泛应用的两类直接方法，即共轭方向法和不计算导数的拟牛顿法．共轭方向法在 Powell(1964) 的工作后，几乎无本质的进展．修正后的 Powell 方法仍是目前最好的直接方法之一．有不少数值结果表明，差分拟牛顿法可能比共轭方向法更好．但也有人认为差分拟牛顿法是利用十分靠近的 $n+1$ 个点来做逼近子问题，其基本思想并不十分好．无导数拟牛顿法利用 n 次小迭代来推导基于函数值的拟牛顿条件．但目前关于它的收敛性还没有很好的结果．

§5.1 交替方向法

最早的也是最简单的直接方法是交替方向法 (alternating directions)，它利用 n 个坐标轴方向进行一维搜索．

算法 5.1.1

步 1　$x_1\in\mathbf{R}^n$ 给出，$\varepsilon\geqslant 0$，$k:=1$．

步 2　$i=1$，$x_k^{(1)}=x_k$；

步 3　求 $\alpha_k^{(i)}$，使得

$$f(x_k^{(i)}+\alpha_k^{(i)}e_i)=\min_{\alpha\in\mathbf{R}^1} f(x_k^{(i)}+\alpha e_i), \tag{5.1.1}$$

令

$$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)}e_i \tag{5.1.2}$$

$i:=i+1$; 如果 $i\leqslant n$, 则转步 3;

步 4 $x_{k+1}=x_k^{(n+1)}$,

如果 $\|x_{k+1}-x_k\|\leqslant\varepsilon$, 则停;

$k:=k+1$, 转步 2.

在算法的步 3 中, e_i 代表第 i 个坐标轴方向. 算法 5.1.1 程序简单, 而且在收敛的情况下必收敛于稳定点.

定理 5.1.2 设 $f(x)$ 连续, 算法 5.1.1 产生的点列 x_k 收敛于 x^*, 则 x^* 必满足

$$f(x^*)=\min_{\alpha\in\mathbb{R}} f(x^*+\alpha e_i) \tag{5.1.3}$$

对一切 $i=1, 2, \cdots, n$ 均成立. 如果 $f(x)$ 在 x^* 处可微, 则 x^* 是 $f(x)$ 的稳定点.

证明 假设定理不真, 则有 x_k 收敛于 x^*, 并且存在 i, 使得 (5.1.3) 不成立. 不失一般性, 假定

$$f(x^*)>\min_{\alpha\in\mathbb{R}^1} f(x^*+\alpha e_1), \tag{5.1.4}$$

则存在 α^*, 使得

$$f(x^*)-f(x^*+\alpha^* e_1)=\delta>0. \tag{5.1.5}$$

由于 x_k 收敛于 x^*, 故对充分大的 k, 都有

$$f(x_k)-f(x_k+\alpha^* e_1)\geqslant\frac{1}{2}\delta. \tag{5.1.6}$$

从而

$$f(x_k)-f(x_{k+1})\geqslant f(x_k)-f(x_k^{(2)})$$

$$\geqslant f(x_k)-f(x_k+\alpha^* e_1)\geqslant\frac{1}{2}\delta \tag{5.1.7}$$

对所有充分大的 k 均成立. 由于 $x_k\to x^*$, 故应该有

$$\lim_{k\to\infty}[f(x_k)-f(x_{k+1})]=0, \tag{5.1.8}$$

这与 (5.1.7) 矛盾. 此矛盾说明定理成立. ∎

定理 5.1.2 的结论是基于点列收敛建立起来的. 但是, 美中不足的是交替方向法可能不收敛, 且任何极限点都不是 $f(x)$ 的稳

定点.

例 5.1.3　极小化函数

$$f(x, y)=\max\{|x|, |y|\}. \tag{5.1.9}$$

我们取初值$(x_1, y_1)=(-1, -1)$,不难发现沿着 e_1 方向搜索任何 $\alpha\in[0, 2]$都是解. 取 $\alpha=2$, 则得到$(x_1^{(2)}, y_1^{(2)})=(1, -1)$. 这时沿着 e_2 方向搜索,我们可取步长 $\alpha=2$, 得到$(x_2, y_2)=(1, 1)$. 类似进行下去,算法 5.1.1 产生的点列为

$$(x_i, y_i)=(-1)^i(1, 1). \tag{5.1.10}$$

而函数$f(x, y)$只有唯一的极小点$(0, 0)$.

算法 5.1.1 的另一个缺点是它和最速下降法一样可能出现 Zigzag(即拐来拐去). 从而收敛十分慢. 所以, Hooke 和 Jeeves (1961)提出模式搜索 (Pattern search)方法, 其基本思想是利用每 n 次交替方向搜索后找到的新点估计出一个方向, 进行一模式搜索步.

算法 5.1.4 (**Hooke 和 Jeeves 方法**)

步 1　$x_1\in\mathbf{R}^n$ 给出, $\varepsilon\geqslant 0$, $k:=1$.

步 2　$i=1$; 如果 $k=1$, 则 $x^{(1)}=x_1$ 以及转步 3;

求 $\alpha_k^{(0)}$ 使得

$$f(x_k+\alpha_k^{(0)}d_k)=\min_{\alpha\in\mathbf{R}}f(x_k+\alpha d_k), \tag{5.1.11}$$

$$x_k^{(1)}=x_k+\alpha_k^{(0)}d_k.$$

步 3　求 $\alpha_k^{(i)}$, 使得(5.1.1)成立, 取

$$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)}e_i \tag{5.1.12}$$

$i:=i+1$, 如果 $i\leqslant n$, 则转步 3;

步 4　$x_{k+1}=x_k^{(n+1)}$, $d_{k+1}=x_{k+1}-x_k$;

如果 $\|d_{k+1}\|\leqslant\varepsilon$, 则停;

$k:=k+1$, 转步 2.

显然, 对于算法 5.1.4, 定理 5.1.2 仍成立. 在一般情况下, 算法 5.1.4 比算法 5.1.1 收敛得快, 但美中不足的是: 算法 5.1.4 并不能保证收敛.

另一个利用交替方向的方法是 Rosenbrock 方法, 也称为转

轴(rotating coordinate)法，它最早是由 Rosenbrock(1960)提出的．这方法每次迭代用一组正交方向进行交替搜索．这个方法后经 Davis、Swann 和 Campey 修改(见 Swann， 1964， 1972)．Palmer(1969)和 Powell(1968)分别讨论了如何比较经济地计算每次迭代中需要的 n 个正交方向．

算法 5.1.5 (Rosenbrock 方法)

步 1　给出 $x_1\in\mathbf{R}^n$, n 个正交方向

$$d_k^{(1)}, \cdots, d_k^{(n)};$$

$$\varepsilon\geqslant 0,\ k:=1;$$

步 2　$i=1$, $x_k^{(1)}=x_k$;

步 3　求 $\alpha_k^{(i)}$, 使得

$$f(x_k^{(i)}+\alpha_k^{(i)}d_k^{(i)})=\min_{\alpha\in\mathbf{R}}f(x_k^{(i)}+\alpha d_k^{(i)}) \tag{5.1.13}$$

令

$$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)}d_k^{(i)} \tag{5.1.14}$$

$i:=i+1$, 如果 $i\leqslant n$, 则转步 3;

步 4　$x_{k+1}=x_k^{(n+1)}$;

如果 $\|x_{k+1}-x_k\|\leqslant\varepsilon$, 则停;

构造新的正交方向 $d_{k+1}^{(i)}(i=1, \cdots, n)$;

$k:=k+1$; 转步 2.

初始的正交方向可取为 $e_i(i=1, \cdots, n)$， 以后每次迭代后，可定义

$$\bar{d}_k^{(i)}=\begin{cases}d_k^{(i)}, & \text{如果 } \alpha_k^{(i)}=0;\\ x_{k+1}-x_k^{(i)}, & \text{如果 } \alpha_k^{(i)}\neq 0.\end{cases} \tag{5.1.15}$$

然后 $d_{k+1}^{(i)}(i=1, 2, \cdots, n)$即可用 Gram-Schinidt 方法将 $\bar{d}_k^{(i)}(i=1, \cdots, n)$正交化而得到．我们有

$$d_{k+1}^{(1)}=\bar{d}_k^{(1)}/\|\bar{d}_k^{(1)}\|_2 \tag{5.1.16}$$

$$d_{k+1}^{(i)}=\bar{d}_k^{(i)}-\sum_{j=1}^{i-1}\bar{d}_k^{(i)T}d_{k+1}^{(j)}d_{k+1}^{(j)}/\rho_k^{(i)} \tag{5.1.17}$$

其中，$i=2, \cdots, n$．$\rho_k^{(i)}$ 是如此选取，使得

$$\|d_{k+1}^{(i)}\|_2=1. \tag{5.1.18}$$

不难看到，Rosenbrock 方法每次迭代用到的 n 个方向是相互正交的，从而它和算法 5.1.1 一样，如果收敛，则必收敛于稳定点。另外，由 Bazaraa 和 Shetty(1979)可知，只要函数 $f(x)$在任何一维子空间上的极小点都是唯一的，且每次迭代的 n 个方向是一致无关的，则点列 x_k 的任何聚点都是 $f(x)$的稳定点. 在这里说方向 $d_k^{(1)}, \cdots, d_k^{(n)}$ 一致无关，是指矩阵

$$U_k=[d_k^{(1)}/\|d_k^{(1)}\|_2, \cdots, d_k^{(n)}/\|d_k\|_2] \tag{5.1.19}$$

的逆存在，且存在常数 M，使得

$$\|U_k^{-1}\|_2\leqslant M \tag{5.1.20}$$

对一切 k 都成立. 对于 Rosenbrock 方法，U_k 是一正交矩阵，于是(5.1.20)显然成立. 从而我们有如下结果：

定理 5.1.6 设 $f(x)$ 连续可微且严格凸，由算法 5.1.5 产生的点列有界，则 x_k 必收敛于 $f(x)$ 的唯一极小点 x^*.

证明 设 $\bar{x}$ 是 x_k 的任何聚点. 由于 U_k 是正交阵，必存在正交向量 $\bar{d}^{(1)}, \cdots, \bar{d}^{(n)}$，使得

$$f(\bar{x})=\min_{\alpha\in\mathbf{R}^1} f(\bar{x}+\alpha\bar{d}^{(i)}). \tag{5.1.21}$$

于是不难证明

$$\bar{d}^{(i)T}\nabla f(\bar{x})=0. \tag{5.1.22}$$

因为 $\bar{d}^{(i)}(i=1, \cdots, n)$线性无关，从而 $\nabla f(\bar{x})=0$. 所以 $\bar{x}=x^*$.故知 $x_k\to x^*$. ▮

§5.2 单纯形法

单纯形法(simplex method)是另一个较早的直接方法. 它最先由 Spendley, Hext 和 Himsworth(1962)提出，后来由 Nelder 和 Mead(1965)改进. 单纯形(simplex)是指 $\mathbf{R}^n$ 中的 $n+1$ 个点为顶点的凸包. 单纯形法的基本思想是：在每次迭代时利用已有的单纯形去寻找一个函数值更小的点. 如果得到这样一个更好的点，则用这个新点作为一个顶点构造新的单纯形. 否则的话，将已有单纯形缩小重复迭代. 值得注意的是：解无约束优化的单纯形法和解线

性规划的单纯形方法是完全不一样的，尽管两者的名字一模一样。

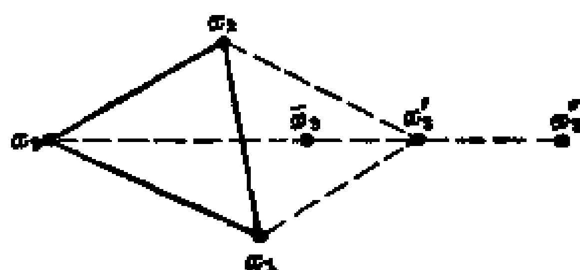


图 5.2.1

下面看简单情形 $n=2$：设给出点 $x_i(i=1,2,3)\in\mathbf{R}^2$，$f(x_3)\geqslant f(x_2)\geqslant f(x_1)$。考虑由 x_1、x_2 和 x_3 组成的单纯形。为了方便叙述，我们特意将 x_1 和 x_2 的连线安排在垂直方向。由于在线段 x_1x_2 的左边的点 x_3 的函数 $f(x)$ 的值较大，所以有理由在 x_1 和 x_2 连线的右边去找一个可能使 $f(x)$ 下降的点。于是我们可试 x_3 的反射点，即

$$x_3'=x_1+x_2-x_3, \tag{5.2.1}$$

如果 $f(x_3')<f(x_1)$，我们得到一个较好的点，这时可试着沿 x_3x_3' 的方向外推寻找一个更好的点。即看是否有

$$f(x_3'')<f(x_3'). \tag{5.2.2}$$

其中

$$x_3''=\frac{x_1+x_2}{2}+\gamma\left(x_3'-\frac{x_1+x_2}{2}\right), \tag{5.2.3}$$

$\gamma>1$ 是一常数。如果 $f(x_3')\geqslant f(x_2)$，由于对称性，不妨假设 $f(x_3')\leqslant f(x_3)$。这时，我们可缩小步长，得到 $\bar{x}_3=\dfrac{x_1+x_2}{2}+\beta\left(x_3'-\dfrac{x_1+x_2}{2}\right)$，其中 $\beta\in(0,1)$。如果 $f(\bar{x}_3)\leqslant f(x_3')$，则用 $\bar{x}_3$ 代替 x_3 组成新的单纯形，否则，我们缩小单纯形，即令

$$x_2:=x_1+\frac{1}{2}(x_2-x_1), \tag{5.2.4}$$

$$x_3:=x_1+\frac{1}{2}(x_3'-x_1). \tag{5.2.5}$$

最后一种情况是 $f(x_1)\leqslant f(x_3')<f(x_2)$。此时，我们可用点 x_1、x_2、x_3' 组成新的单纯形进行下一次迭代。

对一般 $n\geqslant2$，单纯形法的原理和上面讨论的一样，每次迭代时有 $n+1$ 个点 x_1，…，x_{n+1}；设下标 i_h 和 i_l 使得

$$f(x_{i_h})=\max_{1\le j\le n+1} f(x_j), \tag{5.2.6}$$

$$f(x_{i_l})=\min_{1\le j\le n+1} f(x_j). \tag{5.2.7}$$

然后，为了计算反射点，我们需要计算

$$\bar{x}=\frac{1}{n}\sum_{\substack{j=1\\ j\ne i_h}}^{n+1} x_j. \tag{5.2.8}$$

单纯形的终止判别条件是

$$\left[\frac{1}{n+1}\sum_{i=1}^{n+1}\left[f(x_i)-f(\bar{x})\right]^2\right]^{1/2}\le\varepsilon, \tag{5.2.9}$$

其中 $\varepsilon\ge 0$ 是给定的误差许可(见 Nelder 和 Mead, (1965)). 由于判别条件(5.2.9)需要计算函数 $f(x)$ 在 $\bar{x}$ 处的值，我们认为用

$$\left[\frac{1}{n}\sum_{\substack{j=1\\ j\ne i_l}}^{n+1}\left[f(x_j)-f(x_{i_l})\right]^2\right]^{\frac{1}{2}}\le\varepsilon \tag{5.2.10}$$

作为判别条件可能比(5.2.9)更好. 由于对任何 $n+1$ 个点 x_i ($i=1, \cdots, n+1$)的函数值相等时均满足(5.2.10)，所以用(5.2.10)时必须要求 $\sum_{i=1}^{n+1}\|x_i-x_{i_l}\|_2^2$ 也非常小. 下面给出单纯形法.

算法 5.2.1 (单纯形法)

步 1　给出 $x_1, \cdots, x_{n+1}\in\mathbf{R}^n$, $\alpha:=1$, $\beta=0.5$, $\gamma=2$;

步 2　由(5.2.6)～(5.2.8)计算 i_h, i_l 和 $\bar{x}$;

$$x_r=\bar{x}+\alpha(\bar{x}-x_{i_h}); \tag{5.2.11}$$

步 3　如果 $f(x_r)<f(x_{i_l})$，则转步 4

$$x_{i_h}=\begin{cases}x_r, & \text{如果 } f(x_r)\le f(x_r+\gamma(x_r-\bar{x})),\\ x_r+\gamma(x_r-\bar{x}), & \text{反之};\end{cases}$$

转步 6;

步 4　如果 $f(x_r)\ge\max\limits_{\substack{1\le j\le n+1\\ j\ne i_h}} f(x_j)$，则转步 5;

$$x_{i_h}:=x_r$$

转步 6;

步 5　求 x'_{i_h}，使得

$$f(x'_{i_h})=\min\{f(x_{i_h}), f(x_r)\}.$$

计算 $x_c=\bar{x}+\beta(x'_h-\bar{x})$,

如果 $f(x_c)<\max\limits_{\substack{1\le j\le n+1\\ j\ne i_h}} f(x_j)$, 则 $x_{i_h}=x_c$.

否则, 令 $x_i=\dfrac{1}{2}(x_i+x_{i_l})(i=1, \cdots, n+1)$;

步 6 如果(5.2.9)成立, 则停; 转步 2.

单纯形法并没有很好的理论性质. 即使它收敛, 它的收敛速度显然是线性的. 但它具有简单实用的优点. 大量的计算表明单纯形方法是一个十分可靠的方法, 特别是它能处理函数值变化剧烈的函数.

§5.3 共轭方向法

我们在第三章曾看到利用共轭方向可得到二次终止算法. 共轭方向法是利用求平行子空间的极小值的方式来产生共轭方向, 然后利用共轭方向进行搜索.

首先有如下定理:

定理 5.3.1 设 $k<n$, 且 $d_1, \cdots, d_k$ 线性无关, 设

$$S_1=\left\{z \mid z=v_1+\sum_{i=1}^{k}\alpha_i d_i\right\}, \tag{5.3.1}$$

$$S_2=\left\{z \mid z=v_2+\sum_{i=1}^{k}\alpha_i d_i\right\} \tag{5.3.2}$$

是 $\mathbf{R}^n$ 中两个不同的子空间, 设 $z_1\in S_1$ 和 $z_2\in S_2$ 分别是二次凸函数

$$f(x)=g^T x+\frac{1}{2}x^T H x \tag{5.3.3}$$

在 S_1 和 S_2 上的最小点, 则 z_2-z_1 与 $d_1, \cdots, d_k$ 是 H- 共轭的.

证明 显然有

$$d_i^T\nabla f(z_1)=d_i^T\nabla f(z_2)=0(i=1, \cdots, k). \tag{5.3.4}$$

于是

$$d_i^T H(z_2-z_1)=d_i^T[\nabla f(z_2)-\nabla f(z_1)]=0. \qquad (5.3.5)$$

所以定理成立. ∎

共轭方向法的基本思想正是利用定理5.3.1来构造共轭方向的，从而使算法具有二次终止性. 最早的共轭方向法由Smith(1962)提出. 求解二次函数极小的Smith方法可写成如下形式：

算法 5.3.2

步1　给出 $d_1, \cdots, d_n$ 线性无关；$\beta_1, \cdots, \beta_n>0$，

　　给出初值 x_1；

步2　求 α_1：使得 $f(x_1+\alpha d_1)$ 达到最小；

　　$x_2=x_1+\alpha_1 d_1$，$i=2$.

步3　$v^{(1)}=x_i+\beta_i d_i$

　　$j:=1$；

步4　求 $\bar{\alpha}^{(j)}$，使得

$$f(v^{(j)}+\bar{\alpha}^{(j)}d_j)=\min_{\alpha\in\mathbf{R}} f(v^{(j)}+\alpha d_j), \qquad (5.3.6)$$

$$v^{(j+1)}=v^{(j)}+\bar{\alpha}^{(j)}d_j \qquad (5.3.7)$$

　　如果 $j<i$，则转步4；

　　令

$$d_i=v^{(i+1)}-x_i. \qquad (5.3.8)$$

步5　求 α_i 使得 $f(x_i+\alpha d_i)$ 达到最小；

　　$x_{i+1}=x_i+\alpha_i d_i$；

　　$i:=i+1$；如果 $i<n$，则转步3.

利用定理5.3.1即可知算法5.3.2求得的 x_{n+1} 必是函数 $f(x)=g^Tx+\frac{1}{2}x^THx$ 的唯一的极小点 $-H^{-1}g$.

Smith 方法虽然是针对二次函数提出来的，但它也可用来解一般非线性无约束优化问题. Smith (1962) 建议反复调用算法5.3.2，即把 n 次迭代当成一次大的迭代. 但 Fletcher(1965) 发现这种做法具体计算表现并不好. 一个显然的缺点是 Smith 方法对 n 个方向 $d_1, \cdots, d_n$ 的依赖不均. 从算法 5.3.2 可看出：d_1 在每次迭代时都用到，而 d_n 只有在最后一次迭代时才用到.

Powell(1964)提出一个同等对待所有 $d_1, \cdots, d_n$ 的共轭方向法，此方法仍具有二次终止性.

算法 5.3.3

步 1 给出 $d_1, \cdots, d_n$ 线性无关，给出 $x_1 \in \mathbf{R}^n$. $k := 1$;

步 2 $x_k^{(1)} = x_k$, $i = 1$;

步 3 求 $\alpha_k^{(i)}$ 使得

$$f(x_k^{(i)} + \alpha_k^{(i)} d_i) = \min_{\alpha \in \mathbf{R}^1} f(x_k^{(i)} + \alpha d_i), \qquad (5.3.9)$$

令 $x_k^{(i+1)} = x_k^{(i)} + \alpha_k^{(i)} d_i$; $i = i + 1$,

如果 $i \leqslant n$，则转步 3;

步 4 $d_j := d_{j+1} (j = 1, 2, \cdots, n-1)$.

$d_n := x_k^{(n+1)} - x_k$,

求 α_k^* 使得

$$f(x_k + \alpha_k^* d_n) = \min_{\alpha \in \mathbf{R}} f(x_k + \alpha d_n), \qquad (5.3.10)$$

$x_{k+1} = x_k + \alpha_k^* d_n$,

$k := k + 1$; 转步 2.

对于二次凸函数，Powell 方法大约需要比 Smith 方法多一倍精确搜索才会终止. 但是，由于 Powell 方法中各搜索方向 $d_1, \cdots, d_n$ 是平等对待的，方法可理解为在每次迭代都有 n 个可能共轭(或者说近似共轭)的方向. 实际计算表明，Powell 方法对于一般非线性函数 $f(x)$ 比 Smith 方法更有效.

但是，Powell 方法的一个弱点是随着迭代次数的增加，方向 $d_1, \cdots, d_n$ 很容易趋于线性相关. 而且假定在某次迭代时 $d_1, \cdots, d_n$ 线性相关，不难发现对所有的以后迭代点 x_k 均会落在一个子空间内，故只能找到这子空间上的极小点.

Powell 提出了一个修改方案，其思想是基于下面结果：

引理 5.3.4 设矩阵 H 正定，$d_i^T H d_i = 1 (i = 1, \cdots, n)$，则当且仅当 $d_1, \cdots, d_n$ 是 H-共轭时行列式

$$|\det([d_1, \cdots, d_n])| \qquad (5.3.11)$$

达到最大.

证明 记 $D=(d_1, \cdots, d_n)$，则有

$$[\det(D)]^2\det(H)=\det(D^THD)$$
$$\leqslant\frac{1}{n}\operatorname{Tr}(D^THD)=1. \qquad (5.3.12)$$

从而可知(5.3.11)达到最大当且仅当

$$\det(D^THD)=\frac{1}{n}\operatorname{Tr}(D^THD). \qquad (5.3.13)$$

(5.3.13)成立当且仅当 $D^THD=I$，所以引理为真. ∎

现在考虑将 $x_k^{(n+1)}-x_n$ 替换 d_i，这时我们记第 k 次迭代的 n 个方向为 $d_k^{(1)}, \cdots, d_k^{(n)}$，以及 $D_k=[d_k^{(1)}, \cdots, d_k^{(n)}]$，则有

$$D_{k+1}=\left[d_k^{(1)}, \cdots, d_k^{(i-1)}, \sum_{j=1}^{n}\alpha_k^{(j)}d_k^{(j)}, d_k^{(i+1)}, \cdots, d_k^{(n)}\right]. \qquad (5.3.14)$$

由引理 5.3.4，我们希望

$$\frac{[\det(D_{k+1})]^2}{\prod_{j=1}^{n}d_{k+1}^{(j)T}Hd_{k+1}^{(j)T}}=\rho_{k+1} \qquad (5.3.15)$$

达到最大. 利用(5.3.14)可知

$$\rho_{k+1}=\rho_k\cdot[\alpha_k^{(i)}]^2d_k^{(i)T}Hd_k^{(i)}/(x_k^{(n+1)}-x_k)^TH(x_k^{(n+1)}-x_k). \qquad (5.3.16)$$

由于我们假定 $f(x)$ 是凸二次函数(5.3.3)，我们有

$$f(x_k^{(i)})-f(x_k^{(i+1)})=\frac{1}{2}[\alpha_k^{(i)}]^2(d_k^{(i)})^THd_k^{(i)}, \qquad (5.3.17)$$

$$f(x_k)-f(x_{k+1})=\frac{(\alpha_k^*)^2}{2}(x_k^{(n+1)}-x_k)^TH(x_k^{(n+1)}-x_k). \qquad (5.3.18)$$

于是从(5.3.16)～(5.3.18)知

$$\rho_{k+1}=(\alpha_k^*)^2\frac{f(x_k^{(i)})-f(x_k^{(i+1)})}{f(x_k)-f(x_{k+1})}\rho_k. \qquad (5.3.19)$$

由(5.3.19)即知，我们应当选取 i，使得

$$f(x_k^{(i)})-f(x_k^{(i+1)}) \qquad (5.3.20)$$

达到最大. 为了使 $d_k^{(1)}, \cdots, d_k^{(n)}$ 的共轭度不降低，我们要求 $\rho_{k+1}\geqslant\rho_k$. 如果

$$(\alpha_k^*)^2 \frac{\max\limits_{1<j<n}[f(x_k^{(j)})-f(x_k^{(j+1)})]}{f(x_k)-f(x_{k+1})}<1, \qquad (5.3.21)$$

则显然由(5.3.19)定义的 ρ_{k+1} 必小于 ρ_k。在这种情形下，我们就不替换任何 $d_k^{(i)}$，而取 $D_{k+1}=D_k$。修改后的 Powell 方法可写成下面的形式：

算法 5.3.5 (Modified Powell method)

步 1 给出线性无关的 $d_1, \cdots, d_n$,
$x_1 \in \mathbf{R}^n$, $k:=1$.

步 2 $x_k^{(1)}=x_k$;
对 $i=1, \cdots, n$, 作

$$\alpha_k^{(i)}: \min f(x_k^{(i)}+\alpha d_i);$$

$$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)} d_i.$$

步 3 $d_{n+1}=x_k^{(n+1)}-x_k$
$\alpha_k^*: \min f(x_k+\alpha d_{n+1})$
$x_{k+1}=x_k+\alpha_k^* d_{n+1}$.

步 4 求 i: $\max\limits_{1<j<n}[f(x_k^{(j)})-f(x_k^{(j+1)})]$.
如果

$$[\alpha_k^*]^2<\frac{f(x_k)-f(x_{k+1})}{f(x_k^{(i)})-f(x_k^{(i+1)})}, \qquad (5.3.22)$$

则转步 5;
对 $j=i, \cdots, n$, 作 $d_j:=d_{j+1}$,

步 5 $k:=k+1$; 转步 2.

Powell 的修正方法是基于 $f(x)$ 是二次函数，但对于一般非线性函数 $f(x)$，(5.3.17)与(5.3.18)并不成立，从而条件(5.3.22)并不能保证随着迭代的增加 $d_k^{(1)}, \cdots, d_k^{(n)}$ 的共轭性越好. 由于对非线性函数 H 并不知道，所以 Zangwill(1967a) 建议每次迭代要求

$$|\det(D_{k+1})|^2 \Big/ \prod_{j=1}^{n} \|d_{k+1}^{(i)}\|_2^2 \qquad (5.3.23)$$

尽可能大，这可理解取 $H=I$. 假定 $d_k^{(i)}$ $(i=1, \cdots, n)$ 都是单位长

度向量，令 $d_{k+1}^{(i)}=\sum_{j=1}^{n}\alpha_k^{(j)}d_k^{(j)}\Big/\left\|\sum_{j=1}^{n}\alpha_k^{(j)}d_k^{(j)}\right\|_2$ 且 $d_{k+1}^{(j)}=d_k^{(j)}$ $(j=1,\cdots,n,\ j\neq i)$. 则有

$$\det(D_{k+1})=\frac{|\alpha_k^{(i)}|}{\left\|\sum_{j=1}^{n}\alpha_k^{(j)}d_k^{(j)}\right\|_2}\det(D_k). \tag{5.3.24}$$

利用(5.3.24)即可导出 Zangwill 的修正 Powell 方法如下：

算法 5.3.6　(Zangwill's Method)

步 1　给出 $d_1,\cdots,d_n$ 且 $\Delta=|\det(d_1,\cdots,d_n)|\geqslant\delta>0$.

　　$x_1\in\mathbf{R}^n$, $k:=1$.

步 2　$x_k^{(1)}=x_k$;

　　对 $i=1,\cdots,n$, 作

　　$\alpha_k^{(i)}:\min f(x_k^{(i)}+\alpha d_i)$;

　　$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)}d_i$.

步 3　$d_{n+1}=(x_k^{(n+1)}-x_k)/\|x_k^{(n+1)}-x_k\|_2$;

　　$\alpha_k^*:\min f(x_k+\alpha d_{n+1})$;

　　$x_{k+1}=x_k+\alpha_k^*d_{n+1}$.

步 4　求 $i:\max\limits_{1\leqslant j\leqslant n}|\alpha_k^{(j)}|$

　　如果

$$\frac{|\alpha_k^{(i)}|\,\Delta}{\|x_k^{(n+1)}-x_k\|_2}<\delta, \tag{5.3.25}$$

　　则转步 5;

　　$d_i:=d_{n+1}$;

$$\Delta:=\Delta|\alpha_k^{(i)}|/\|x_k^{(n+1)}-x_k\|_2. \tag{5.3.26}$$

步 5　$k:=k+1$, 转步 2.

不难看到算法 5.3.6 和算法 5.3.4 的差别在于判别条件(5.3.25)与(5.3.22)的不同. 由于算法 5.3.6 保证

$$|\det(D_k)|\geqslant\delta \tag{5.3.27}$$

对一切 k 均成立. 所以 $d_k^{(1)},\cdots,d_k^{(n)}$ 一定线性无关. 由此，我们有收敛性定理如下：

定理 5.3.7　设 $f(x)$ 连续可微且严格凸，如果算法 5.3.6 产

生的点列$\{x_k\}$有界，则x_k收敛于$f(x)$唯一的极小点.

证明 设$\bar{x}$是$\{x_k\}$的任一聚点. 不失一般性，我们可假定$d_{k_i}^{(1)}\to d^{(1)}$, $\alpha_{k_i}^{(1)}\to\alpha^{(1)}$. 显然有$\|d^{(1)}\|_2=1$. 由于算法是下降的，故有

$$f(\bar{x})\geqslant f(\bar{x}+\alpha^{(1)}d^{(1)}), \tag{5.3.28}$$

如果$f(\bar{x})>f(\bar{x}+\alpha^{(1)}d^{(1)})$，则存在$\eta>0$使得

$$f(x_{k_i})\geqslant f(x_{k_i}+\alpha_{k_i}^{(1)}d_{k_i}^{(1)})+\eta\geqslant f(x_{k_{i+1}})+\eta. \tag{5.3.29}$$

于是$f(x_k)\to-\infty$. 这与$\{x_k\}$有界相矛盾. 所以有

$$f(\bar{x})=f(\bar{x}+\alpha^{(1)}d^{(1)}). \tag{5.3.30}$$

设$\alpha_*^{(1)}$是$\min f(\bar{x}+\alpha d^{(1)})$的解. 由于$f(x)$严格凸，如果$\alpha_*^{(1)}\neq 0$，则有$\bar{\alpha}^{(1)}\in\mathbb{R}^1$使得

$$f(\bar{x}+\bar{\alpha}^{(1)}d^{(1)})<f(\bar{x}). \tag{5.3.31}$$

于是，对充分大的i，有

$$\begin{aligned}f(x_{k_i})-f(x_{k_{i+1}})&\geqslant f(x_{k_i})-f(x_{k_i}+\alpha_{k_i}^{(1)}d_{k_i}^{(1)})\\&\geqslant f(x_{k_i})-f(x_{k_i}+\bar{\alpha}^{(1)}d_{k_i}^{(1)})\\&\geqslant\frac{1}{2}\left[f(\bar{x})-f(\bar{x}+\bar{\alpha}^{(1)}d^{(1)})\right]>0.\end{aligned} \tag{5.3.32}$$

这与$\{x_k\}$有界相矛盾. 此矛盾说明$\alpha_*^{(1)}=0$. 于是必有$\alpha^{(1)}=0$. 从而$x_{k_i}^{(2)}\to\bar{x}$. 依此我们可假定

$$d_{k_i}^{(j)}\to d^{(j)}(j=1,\ 2,\ \cdots,\ n), \tag{5.3.33}$$

$$\alpha_{k_i}^{(j)}\to\alpha^{(j)}(j=1,\ 2,\ \cdots,\ n). \tag{5.3.34}$$

同上可证$\alpha^{(j)}=0(j=1,\ \cdots,\ n)$且

$$f(\bar{x})=\min_{\alpha\in\mathbb{R}}f(\bar{x}+\alpha d^{(j)}). \tag{5.3.35}$$

由(5.3.35)即知

$$[d^{(j)}]^T\nabla f(\bar{x})=0(j=1,\ \cdots,\ n). \tag{5.3.36}$$

由于$|\det(D_k)|\geqslant\delta$，故必有

$$|\det(d^{(1)},\ \cdots,\ d^{(n)})|\geqslant\delta>0. \tag{5.3.37}$$

从(5.3.36)和(5.3.37)知

$$\nabla f(\bar{x})=0. \tag{5.3.38}$$

所以$\bar{x}$是$f(x)$的唯一极小点x^*. 由于算法是下降算法，从而有

$$\lim_{k\to\infty} f(x_k^{(j)})=f(x^*) \tag{5.3.39}$$

对 $j=1, 2, \cdots, n$ 都成立. 由于 $f(x)$ 严格凸, 所以 $x_k^{(j)}\to x^*(j=1, \cdots, n)$. 定理得证. ▮

§5.4 差分拟牛顿法

差分拟牛顿法是用差商代替导数的拟牛顿法. 这类方法是共轭方向法之外的另一类有效的直接方法.

差分拟牛顿法最早是由 Stewart (1967) 提出的, 他研究了将 DFP 方法差分化. 后来 Gill 和 Murray (1972) 考虑了差分形式下的 Broyden 族方法. 差分拟牛顿法和拟牛顿法的区间在于每次迭代时 $g_k=\nabla f(x_k)$ 用差商

$$\tilde{g}_k=\begin{bmatrix} \dfrac{f(x_k+h_k e_1)-f(x_k)}{h_k} \\ \vdots \\ \dfrac{f(x_k+h_k e_n)-f(x_k)}{h_k} \end{bmatrix} \tag{5.4.1}$$

代替. (5.4.1)用到的是向前差分. 为了提高逼近精度, 也可用中间差分, 即取

$$\tilde{g}_k=\begin{bmatrix} \dfrac{f(x_k+h_k e_1)-f(x_k-h_k e_1)}{2h_k} \\ \vdots \\ \dfrac{f(x_k+h_k e_n)-f(x_k-h_k e_n)}{2h_k} \end{bmatrix}. \tag{5.4.2}$$

但中间差分需要多计算 n 个函数值. 差分步长 $h_k>0$ 的选取对算法的影响是十分大的. 差分拟牛顿法有如下形式:

算法 5.4.1 (差分拟牛顿法)

步 1 给出初始值 $x_1\in \mathbf{R}^n$, $H_1\in \mathbf{R}^{n\times n}$ 对称,

利用差商计算 $\tilde{g}_1$, $k:=1$;

步 2 $d_k=-H_k\tilde{g}_k$

线性搜索求 α_k,

$$x_{k+1}=x_k+\alpha_k d_k$$

步 3 利用差商求 $\tilde{g}_{k+1}$.

$$\tilde{y}_k=\tilde{g}_{k+1}-\tilde{g}_k. \tag{5.4.3}$$

用 $s_k=x_{k+1}-x_k$ 和 $\tilde{y}_k$ 修正矩阵 H_{k+1}.

步 4 $k:=k+1$, 转步 2.

差分拟牛顿法不计算导数，所以进行线搜索时只能用仅依赖于函数值的一维极小化方法.

每次迭代时的差商步长在各个方向上可以不等. 即可将(5.4.1)式换成

$$\tilde{g}_k=\begin{pmatrix}\dfrac{f(x_k+h_k^{(1)}e_1)-f(x_k)}{h_k^{(1)}}\\ \vdots\\ \dfrac{f(x_k+h_k^{(n)}e_1)-f(x_k)}{h_k^{(n)}}\end{pmatrix}. \tag{5.4.4}$$

Stewart(1967) 要求差商步长满足一定的误差界. 假定 $\eta>0$ 是一给定的误差界，设 $h_k^{(i)}$ 是在第 k 次迭代在第 i 个坐标轴方向的差商步长，则

$$\tilde{g}_k^{(i)}=\frac{f(x_k+h_k^{(i)}e_i)-f(x_k)}{h_k^{(i)}}. \tag{5.4.5}$$

记 H_k^{-1} 的第 i 个对角线元素为 $\alpha_k^{(i)}$. 在一般情况下 $|\alpha_k^{(i)}|\neq 0$ (对于差分变尺度法，即要求所有 H_k 正定，显然有 $\alpha_k^{(i)}>0$). 取

$$\bar{h}_k^{(i)}=\begin{cases}2\left[\dfrac{f(x_k)}{|\alpha_k^{(i)}|}\right]^{\frac{1}{2}}, \text{如果 } (\tilde{g}_k^{(i)})^2\geqslant|\alpha_k^{(i)}f(x_k)|\eta;\\ 2\left(\dfrac{|f(x_k)\tilde{g}_k^{(i)}|}{[\alpha_k^{(i)}]^2}\right)^{\frac{1}{3}}, \text{反之.}\end{cases} \tag{5.4.6}$$

以及

$$h_k^{(i)}=\begin{cases}\bar{h}_k^{(i)}\left(1-\dfrac{|\alpha_k^{(i)}|\bar{h}_k^{(i)}}{3|\alpha_k^{(i)}|\bar{h}_k^{(i)}+4|\tilde{g}_k^{(i)}|}\right), \text{如果 } (\tilde{g}_k^{(i)})^2\geqslant|\alpha_k^{(i)}f(x_k)|\eta;\\ \bar{h}_k^{(i)}\left(1-\dfrac{2|\tilde{g}_k^{(i)}|}{3|\alpha_k^{(i)}|\bar{h}_k^{(i)}+4|\tilde{g}_k^{(i)}|}\right), \text{反之.}\end{cases} \tag{5.4.7}$$

如果对于给定的精度要求 $\varepsilon>0$，有

$$|\hat{h}_k^{(i)}|<2\varepsilon|\tilde{g}_k^{(i)}|/|\alpha_k^{(i)}|, \tag{5.4.8}$$

则在 $|h_k^{(i)}|\leqslant|\hat{h}_k^{(i)}|$ 时，我们取(5.4.5)作为近似方向导数；否则，我们取

$$\tilde{g}_k^{(i)}=\frac{f(x_k+t\hat{h}_k^{(i)}e_i)-f(x_k)}{t\hat{h}_k^{(i)}}, \tag{5.4.9}$$

其中 $t=\mathrm{sign}(\alpha_k^{(i)})\mathrm{sign}([f(x_k+h_k^{(i)}e_i)-f(x_k)]/h_k^{(i)})$.如果(5.4.8)不满足，则令新的 $h_k^{(i)}$ 为方程

$$\frac{1}{2}|\alpha_k^{(i)}|h^2+|\tilde{g}_k^{(i)}|h-\frac{1}{\varepsilon}|f(x_k)|\eta=0 \tag{5.4.10}$$

的正根，取

$$\tilde{g}_k^{(i)}=\frac{f(x_k+h_k^{(i)}e_i)-f(x_k-h_k^{(i)}e_i)}{2h_k^{(i)}}. \tag{5.4.11}$$

关于差分步长选取的详细讨论，可参阅 Stewart (1967) 与Gill 和 Murray(1972). 差分变尺度法的收敛性分析可见赵小平(1990).

差分拟牛顿法是直接将原有的拟牛顿法“直接化”. 另一种不需计算导数的拟牛顿法，是无导数拟牛顿(quasi-Newton without derivatives)法. 它是基于修正矩阵使其只依赖函数值. 这方法由 Greenstadt(1972a)提出，并在 Greenstadt (1972b)中加以改进.

无导数拟牛顿法的基本思想是利用只依赖于函数值的某种近似于拟牛顿公式的条件.它每次迭代进行 n 次一维搜索.在第 k 次迭代时，假定有在当前迭代点 x_k 的一个近似梯度 $\tilde{g}_k$ 以及一个近似海色阵 B_k，算法第一个一维搜索方向为

$$d_k^{(1)}=-B_k^{-1}\tilde{g}_k. \tag{5.4.12}$$

将其单位长度化后得到

$$\bar{d}_k^{(1)}=d_k^{(1)}/\|d_k^{(1)}\|_2. \tag{5.4.13}$$

对于 $i=2,\cdots,n$，我们取

$$d_k^{(i)}=e_i-\sum_{j=1}^{i-1}\bar{d}_k^{(j)}(\bar{d}_k^{(j)})^TB_ke_i, \tag{5.4.14}$$

$$\bar{d}_k^{(i)}=d_k^{(i)}/\|d_k^{(i)}\|_2. \tag{5.4.15}$$

记 $x_k^{(1)}=x_k$, 对所有的 $i=1,\cdots,n$, 取 $\alpha_k^{(i)}$ 使得

$$\min f(x_k^{(i)}+\alpha\bar{d}_k^{(i)}) \tag{5.4.16}$$

达到极小，且定义

$$x_k^{(i+1)}=x_k^{(i)}+\alpha_k^{(i)}\bar{d}_k^{(i)}. \tag{5.4.17}$$

当 $f(x)$ 是二次凸函数时

$$f(x_k^{(i)})-f(x_k^{(i+1)})=\frac{1}{2}(s_k^{(i)})^T H s_k^{(i)}, \tag{5.4.18}$$

$$(s_k^{(i)})^T[g_k+H(x_k^{(i+1)}-x_k)]=0. \tag{5.4.19}$$

其中 $s_k^{(i)}=x_k^{(i+1)}-x_k^{(i)}$. 于是我们要求 $\hat{g}_k$ 和 B_{k+1}, 使得

$$(s_k^{(i)})^T B_{k+1} s_k^{(i)}=2[f(x_k^{(i)})-f(x_k^{(i+1)})], \tag{5.4.20}$$

$$(s_k^{(i)})^T[\hat{g}_k+B_{k+1}(x_k^{(i+1)}-x_k)]=0 \tag{5.4.21}$$

对一切 $i=1,2,\cdots,n$ 均成立. (5.4.20)与(5.4.21)可理解为基于函数值的拟牛顿条件. 利用变分法可求得

$$\hat{g}_k=\tilde{g}_k+\sum_{i=1}^{n}\theta_k^{(i)}s_k^{(i)}, \tag{5.4.22}$$

$$B_{k+1}=B_k+\frac{1}{2}\sum_{i=1}^{n}\{\theta_k^{(i)}[s_k^{(i)}(x_k^{(i+1)}-x_k)^T+(x_k^{(i+1)}-x_k)(s_k^{(i)})^T] +\eta_k^{(i)}s_k^{(i)}(s_k^{(i)})^T\}, \tag{5.4.23}$$

其中

$$\theta_k^{(i)}=\frac{2[f(x_k^{(i+1)})-f(x_k^{(i)})]-\tilde{g}_k^T s_k^{(i)}-(s_k^{(i)})^T B_k(x_k^{(i)}-x_k)}{\|s_k^{(i)}\|_2^2\left(\nu_k^2+\frac{1}{2}\sum_{j=1}^{i-1}\|s_k^{(j)}\|_2^2\right)} \tag{5.4.24}$$

$$\eta_k^{(i)}=2\left[\frac{2[f(x_k^{(i)})-f(x_k^{(i+1)})]-(s_k^{(i)})^T B_k s_k^{(i)}}{\|s_k^{(i)}\|_2^4}-\theta_k^{(i)}\right]. \tag{5.4.25}$$

$\nu_k>0$ 是一个参数. B_{k+1} 和 $\hat{g}_k$ 求出后，下一次迭代所需要的近似梯度取为

$$\tilde{g}_{k+1}=\hat{g}_k+B_{k+1}(x_k^{(n+1)}-x_k). \tag{5.4.26}$$

令 $x_{k+1}=x_k^{(n+1)}$, 即完成了第 k 次迭代. 无导数拟牛顿算法可写成如下形式:

算法 5.4.2 (无导数拟牛顿法)

步 1　给出 $x_1 \in \mathbf{R}^n$, $B_1 \in \mathbf{R}^{n \times n}$ 对称, 求得近似梯度 $\tilde{g}_1$.

$k:=1$, $\varepsilon \geqslant 0$.

步 2　如果 $\|\tilde{g}_1\|_2 < \varepsilon$, 则停;

$d_k^{(1)} = -B_k^{-1}\tilde{g}_k$;

对 $i=1, 2, \cdots, n$, 作

由(5.4.13)～(5.4.17)求 $\alpha_k^{(i)}$, $x_k^{(i+1)}$;

步 3　由(5.4.22)～(5.4.25)计算 $\hat{g}_k$、B_{k+1};

$x_{k+1} = x_k^{(n+1)}$

$\tilde{g}_{k+1} = \hat{g}_k + B_{k+1}(x_{k+1} - x_k)$

$k:=k+1$, 转步 2.

参数 ν_k 当 $k \to \infty$ 时趋于 0. 条件(5.4.20)与(5.4.21)是基于精确线搜索推导出来的. 在非精确线搜索的情况下, Greenstadt (1972b)用二次插值推导出类似于(5.4.20)与(5.4.21)的条件, 于是利用变分性质同样可导出计算 $\hat{g}_k$ 和 B_{k+1} 的公式. 据作者所知, 并没有人对无导数拟牛顿法的收敛性进行过分析.

第 6 章

二 次 规 划

二次规划是最简单的带约束非线性规划的问题．它是在线性约束下求解一个二次函数的极小值，可写成

$$\min_{x\in\mathbf{R}^n} g^Tx+\frac{1}{2}x^THx=Q(x), \tag{6.0.1}$$

$$\text{s. t.}\quad a_i^Tx=b_i,\ i=1,\ \cdots,\ m_e, \tag{6.0.2}$$

$$a_i^Tx\geqslant b_i,\ i=m_e+1,\ \cdots,\ m, \tag{6.0.3}$$

其中 $m\geqslant m_e\geqslant 0$；$g\in\mathbf{R}^n$，$H\in\mathbf{R}^{n\times n}$ 且对称，$a_i\in\mathbf{R}^n$，$b_i\in\mathbf{R}$ $(i=1,\ \cdots,\ m)$．有时一些约束是上下界形式：

$$l_i\leqslant x_j\leqslant \mu_j\quad (j=1,\ \cdots,\ n), \tag{6.0.4}$$

其中 x_j 是 x 的第 j 个分量，$l_j\leqslant u_j$ 是变量 x_j 的下界和上界．还有的问题要求变量是非负的：

$$x_j\geqslant 0,\ j=1,\ \cdots,\ n. \tag{6.0.5}$$

条件(6.0.5)可写成如下形式：

$$x\geqslant 0. \tag{6.0.6}$$

二次规划是十分特别的一类约束规划问题，所以对它的求解方法也是特殊的．二次规划有着广泛的应用．不少实际问题，如带约束的线性最小二乘，都可表示为二次规划问题．另外，求解非线性约束优化的逐步二次规划方法在每次迭代中都要求解一个二次规划问题．所以有效地求解二次规划将提高逐步二次规划方法的效率．

§6.1 基 本 性 质

在二次规划问题中，目标函数的海色阵和约束函数的雅可比

阵都是常量，所以它的最优性条件有其特殊形式.

由1.2节的结果，我们有以下定理：

定理 6.1.1　设 x^* 是二次规划问题(6.0.1)～(6.0.3)的局部解，则必存在乘子 $\lambda_i^*(i=1, \cdots, m)$，使得

$$g+Hx^*=\sum_{i=1}^{m}\lambda_i^* a_i, \tag{6.1.1}$$

$$\lambda_i^*[a_i^Tx^*-b_i]=0,\ i=m_e+1, \cdots, m; \tag{6.1.2}$$

$$\lambda_i^*\geqslant 0,\ i=m_e+1, \cdots, m. \tag{6.1.3}$$

且对一切非零 d 满足

$$d^Ta_i=0,\ i\in I(x^*)\cup E, \tag{6.1.4}$$

均有

$$d^THd\geqslant 0. \tag{6.1.5}$$

其中 $E=\{1, \cdots, m_e\}$ 以及

$$I(x^*)=\{i \mid a_i^Tx^*=b_i,\ i=m_e+1, \cdots, m\}. \tag{6.1.6}$$

定理 6.1.2　设 x^* 是问题(6.0.1)～(6.0.3)的一个可行点，如果存在乘子 $\lambda_i^*(i=1, \cdots, m)$ 满足(6.1.1)～(6.1.3)，且对一切非零 d，满足

$$d^Ta_i=0,\ i\in E, \tag{6.1.7}$$

$$d^Ta_i\geqslant 0,\ i\in I(x^*), \tag{6.1.8}$$

$$d^Ta_i=0,\ i\in I(x^*)\ 且\ \lambda_i^*>0, \tag{6.1.9}$$

均有

$$d^THd>0. \tag{6.1.10}$$

则 x^* 必是问题(6.0.1)～(6.0.3)的局部严格极小点.

如取 H 是(正定)半正定矩阵，(6.0.1)中的目标函数是(严格)凸函数，这时问题(6.0.1)～(6.0.3)被称为(严格)凸的二次规划问题. 由于二次规划的可行域非空则必为凸集，所以当目标函数是凸函数时，任何 Kuhn-Tucker 点必为二次规划的全局极小点.

定理 6.1.3　设 H 为半正定矩阵，则 x^* 是二次规划问题(6.0.1)～(6.0.3)的全局极小点当且仅当 x^* 是可行点而且存在乘子 $\lambda_i^*(i=1, \cdots, m)$，使得(6.1.1)～(6.1.3)成立.

设 H 是正定矩阵，由定理 6.1.3 可知，求解(6.0.1)～(6.0.3)

等价于求解

$$g+Hx=A\lambda, \tag{6.1.11}$$

$$a_i^T x=b_i,\ i\in E, \tag{6.1.12}$$

$$a_i^T x\geqslant b_i,\ i\in I, \tag{6.1.13}$$

$$\lambda_i[a_i^T x-b_i]=0,\ i\in I, \tag{6.1.14}$$

$$\lambda_i\geqslant 0,\ i\in I. \tag{6.1.15}$$

其中 $I=\{m_e+1,\ \cdots,\ m\}$, $\lambda=(\lambda_1,\ \cdots,\ \lambda_m)$, 以及

$$A=[a_1,\ \cdots,\ a_m]. \tag{6.1.16}$$

记

$$y=A\lambda-g, \tag{6.1.17}$$

$$t_i=a_i^T x-b_i,\ i\in I. \tag{6.1.18}$$

则(6.1.11)～(6.1.15)可写成下列形式:

$$\begin{bmatrix}-b\\ H^{-1}y\end{bmatrix}=\begin{bmatrix}A^T\\ -I\end{bmatrix}(-x)+\begin{bmatrix}0\\ 0\\ t_{m_e+1}\\ \vdots\\ t_m\\ 0\\ \vdots\\ 0\end{bmatrix} \tag{6.1.19}$$

$$A\lambda-y=g, \tag{6.1.20}$$

$$\lambda_i\geqslant 0,\ i\in I, \tag{6.1.21}$$

$$t_i^T\lambda_i=0,\ i\in I, \tag{6.1.22}$$

$$t_i\geqslant 0,\ i\in I. \tag{6.1.23}$$

由定理 6.1.3 可知(6.1.19)～(6.1.23)等价于求解

$$\max b^T\lambda-\frac{1}{2}yH^{-1}y=\bar{Q}(\lambda,\ y), \tag{6.1.24}$$

$$\text{s. t.}\quad A\lambda-y=g, \tag{6.1.25}$$

$$\lambda_i\geqslant 0,\ i\in I. \tag{6.1.26}$$

由于问题(6.1.24)～(6.1.26)与(6.0.1)～(6.0.3)的等价性，我们称 (6.1.24)～(6.1.26) 为 (6.0.1)～(6.0.3) 的对偶(dual)问

题，称(6.0.1)～(6.0.3)为原始(prime)问题. 对偶问题(6.1.24)～(6.1.26)还可化成更简单的形式：

$$\max(b+AH^{-1}g)^T\lambda-\frac{1}{2}\lambda^T(A^TH^{-1}A)\lambda, \quad (6.1.27)$$

$$\text{s.t.} \quad \lambda_i\geqslant 0,\ i\in I. \quad (6.1.28)$$

假定(λ, y)是(6.1.24)～(6.1.26)的可行点，x是(6.0.1)～(6.0.3)的可行点，我们有

$$\begin{aligned}
&Q(x)-\bar{Q}(\lambda, y)\\
&=x^T[A\lambda-y]+\frac{1}{2}x^THx\\
&\quad-\left[\lambda^T[Ax-b]-\sum_{i\in I}\lambda_it_i-\frac{1}{2}yH^{-1}y\right]\\
&=\sum_{i\in I}\lambda_it_i+\frac{1}{2}[x^THx+yH^{-1}y-2x^Ty]. \quad (6.1.29)
\end{aligned}$$

其中t_i由(6.1.18)定义. 由于H正定，不难看出

$$Q(x)\geqslant\bar{Q}(\lambda, y). \quad (6.1.30)$$

(6.1.30)两端相等当且仅当

$$\sum_{i\in I}\lambda_it_i=0 \quad (6.1.31)$$

和

$$x=H^{-1}y. \quad (6.1.32)$$

(6.1.32)显然等价于

$$Hx+g=A\lambda. \quad (6.1.33)$$

从而证明了以下定理：

定理 6.1.4 设H正定，如果原始问题有可行点，则x^*是问题(6.0.1)～(6.0.3)的解当且仅当存在(λ^*, y^*)为问题(6.1.24)～(6.1.26)的解且$x^*=H^{-1}y^*$.

对于原始问题无可行点的情形，我们有以下结果：

定理 6.1.5 设H正定，原始问题无可行点当且仅当对偶问题无界.

证明 如果原始问题有可行点，由(6.1.30)即知对偶问题的目标函数在满足(6.1.25)与(6.1.26)的集合上一致有界.

现在假定原始问题无可行点，于是

$$(a_i^T \quad b_i)\tilde{x}=0,\ i\in E, \tag{6.1.34}$$

$$(a_i^T \quad b_i)\tilde{x}\geqslant 0,\ i\in I, \tag{6.1.35}$$

$$(0 \ \cdots \ 0 \quad 1)\tilde{x}<0. \tag{6.1.36}$$

在 $\tilde{x}\in \mathbf{R}^{n+1}$ 上无解，由 Farkas 引理(引理 1.2.8)即知存在 $\bar{\lambda}_i(i=1,\ \cdots,\ m)$，使得

$$\sum_{i=1}^{m}\bar{\lambda}_i a_i=0, \tag{6.1.37}$$

$$\sum_{i=1}^{m}\bar{\lambda}_i b_i=1, \tag{6.1.38}$$

$$\bar{\lambda}_i\geqslant 0,\ i\in I. \tag{6.1.39}$$

令 $\lambda_i=t\bar{\lambda}_i,\ y=-g$，当 $t\to+\infty$ 时，有

$$\bar{Q}(\lambda,\ y)=t\to+\infty. \tag{6.1.40}$$

所以对偶问题无界．故知定理为真． ▊

原始问题(6.0.1)～(6.0.3)的 Lagrange 函数是

$$L(x,\ \lambda)=Q(x)-\sum_{i=1}^{m}\lambda_i(a_i^T x-b_i). \tag{6.1.41}$$

不难看出，求解 (6.1.11)～(6.1.15) 等价于求 $L(x,\ \lambda)$ 在区域 $\{(x,\ \lambda)\,|\,\lambda_i\geqslant 0,\ i\in I\}$ 上的稳定点．由于 $L(x,\ \lambda)$的海色阵为

$$\nabla^2 L=\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix}, \tag{6.1.42}$$

故有

$$\begin{bmatrix} I & 0 \\ A^T H^{-1} & I \end{bmatrix}\nabla^2 L\begin{bmatrix} I & H^{-1}A \\ 0 & I \end{bmatrix}=\begin{bmatrix} H & 0 \\ 0 & -A^T H^{-1}A \end{bmatrix}. \tag{6.1.43}$$

由于 H 正定，故 $A^T H^{-1}A$ 必为半正定的．从(6.1.43)可知 $\nabla^2 L$ 恰恰有 n 个正特征值，而且它的负特征值的个数正好为 A 的秩．所以 $L(x,\ \lambda)$的稳定点总是一个鞍点(saddle point)．只有在退化情形 $A=0,\ b=0$ 时，$L(x,\ \lambda)=Q(x)$．这时 $L(x,\ \lambda)$的稳定点是 $L(x,\ \lambda)$的极小点．

设$(x^*,\ \lambda^*)$是(6.1.11)～(6.1.15)的解，则对任何(6.0.2)与(6.0.3)的可行点 x，我们有

$$L(x,\ \lambda^*)=Q(x)-\sum_{i=1}^{m}\lambda_i^*(a_i^Tx-b_i)$$

$$=Q(x^*)-\sum_{i=1}^{m}\lambda_i^*(a_i^Tx^*-b_i)$$

$$+Q(x)-Q(x^*)-\sum_{i\in I}\lambda_i^*(a_i^Tx-b_i)$$

$$=L(x^*,\ \lambda^*)+\frac{1}{2}(x-x^*)H(x-x^*)$$

$$\geqslant L(x^*,\ \lambda^*). \tag{6.1.44}$$

另一方面，对任何对偶问题(6.1.27)与(6.1.28)的可行点 λ，不难得到

$$L(x^*,\ \lambda)=Q(x^*)-\sum_{i=1}^{m}\lambda_i(a_i^Tx^*-b_i)$$

$$=L(x^*,\ \lambda^*)-\sum_{i\in I}\lambda_i(a_i^Tx^*-b_i)$$

$$\leqslant L(x^*,\ \lambda^*). \tag{6.1.45}$$

反过来，对任意给出的 $x^*\in\mathbf{R}^n$ 和 $\lambda^*\in\mathbf{R}^m$，如果

$$L(x,\ \lambda^*)\geqslant L(x^*,\ \lambda^*)\geqslant L(x^*,\ \lambda) \tag{6.1.46}$$

对一切满足(6.0.2)与(6.0.3)的 x 和一切满足(6.1.28)的 λ 都成立，则对所有这些 x 和 λ，我们有

$$Q(x)-\sum_{i=1}^{m}\lambda_i^*(a_i^Tx-b_i)\geqslant Q(x^*)-\sum_{i=1}^{m}\lambda_i^*(a_i^Tx^*-b_i), \tag{6.1.47}$$

$$(\lambda-\lambda^*)^T\nabla_\lambda L(x^*,\ \lambda^*)\leqslant 0. \tag{6.1.48}$$

由(6.1.48)即知

$$(\lambda-\lambda^*)^T\nabla_\lambda L(x^*,\ \lambda^*)>0, \tag{6.1.49}$$

$$\lambda_i\geqslant 0,\ i\in I \tag{6.1.50}$$

无解．利用 Farkas 引理(引理 1.2.8)即知 x^* 必满足(6.1.12)与(6.1.13)．由(6.1.49)可知

$$\lambda_i(a_i^Tx^*-b_i)=0,\ i\in I. \tag{6.1.51}$$

假定

$$\lambda_i^*\geqslant 0,\ i\in I. \tag{6.1.52}$$

则由(6.1.51)和(6.1.52)可证

$$\sum_{i=1}^{m}\lambda_i^*(a_i^Tx-b_i)\geqslant\sum_{i=1}^{m}\lambda_i^*(a_i^Tx^*-b_i) \tag{6.1.53}$$

对一切满足(6.0.2)与(6.0.3)的 x 均成立. 从(6.1.47)和(6.1.53)即可推出:

$$Q(x)\geqslant Q(x^*). \tag{6.1.54}$$

于是可知 x 是原始问题(6.0.1)~(6.0.3)的解. 因此我们得到了如下结果:

定理 6.1.6 设 H 正定,则 x^* 是(6.0.1)~(6.0.3)的解当且仅当存在满足(6.1.52)的 $\lambda_i^*(i=1, \cdots, m)$, 使得对一切满足(6.0.2)与(6.0.3)的 x 和对一切满足(6.1.28)的 λ 都有(6.1.46)成立.

定理 6.1.6 说明求凸二次规划的解等价于求它的 Lagrange 函数的鞍点.

二次规划的对偶理论最先由 Dennis(1959)和 Dorn(1960)等人进行研究的. 事实上,对偶理论可推广到凸规划

$$\min_{x\in \mathbf{R}^n} f(x) \tag{6.1.55}$$

$$\text{s. t.}\quad C_i(x)\geqslant 0,\ i=1, \cdots, m). \tag{6.1.56}$$

其中 $f(x)$ 是凸函数, $C_i(x)(i=1, \cdots, m)$ 是凹函数. 关于凸规划的详细讨论可见 Rockafellar(1970).

§6.2 等式约束

在本节中讨论只有等式约束的二次规划问题:

$$\min_{x\in \mathbf{R}^n} g^Tx+\frac{1}{2}x^THx=Q(x). \tag{6.2.1}$$

$$\text{s. t.}\quad A^Tx=b. \tag{6.2.2}$$

其中 $g\in\mathbf{R}^n$, $b\in\mathbf{R}^m$, $A\in\mathbf{R}^{m\times n}$, $H\in\mathbf{R}^{n\times n}$, 且 H 对称. 不失一般性, 可假定秩$(A)=m$.

首先, 我们考虑变量消去(elimination)方法. 假定我们已找到 x 的一个分解 $x=(x_B, x_N)^T$. 其中 $x_B\in\mathbf{R}^m$, $x_N\in\mathbf{R}^{n-m}$, 且对应

的分解 $A=[A_B, A_N]$ 有 A_B 可逆．则知约束条件(6.2.2)等价于

$$A_B^T x_B + A_N^T x_N = b. \tag{6.2.3}$$

由于 A_B^{-1} 存在，故知

$$x_B = A_B^{-T}[b - A_N^T x_N]. \tag{6.2.4}$$

将(6.2.4)代入(6.2.1)，即得到(6.2.1)与(6.2.2)的一个等价形式：

$$\min \hat{g}_N^T x_N + \frac{1}{2} x_N^T \hat{H}_N x_N. \tag{6.2.5}$$

其中

$$\hat{g}_N = g_N - A_N A_B^{-1} g_B + [H_{NB} - A_N A_B^{-1} H_{BB}] A_B^{-T} b, \tag{6.2.6}$$

$$\hat{H}_N = H_{NN} - H_{NB} A_B^{-T} A_N^T - A_N A_B^{-1} H_{BN} + A_N A_B^{-1} H_{BB} A_B^{-T} A_N^T, \tag{6.2.7}$$

且

$$g = \begin{bmatrix} g_B \\ g_N \end{bmatrix}, \tag{6.2.8}$$

$$H = \begin{bmatrix} H_{BB} & H_{BN} \\ H_{NB} & H_{NN} \end{bmatrix} \tag{6.2.9}$$

是对应 $x=(x_B, x_N)^T$ 的分解．

如果 $\hat{H}_N$ 正定，则显然(6.2.5)的解是

$$x_N^* = -\hat{H}_N^{-1} \hat{g}_N. \tag{6.2.10}$$

从(6.2.4)和(6.2.10)可知问题(6.2.1)与(6.2.2)的解为

$$x^* = \begin{bmatrix} x_B^* \\ x_N^* \end{bmatrix} = \begin{bmatrix} A_B^{-T} b \\ 0 \end{bmatrix} + \begin{bmatrix} A_B^{-T} A_N^T \\ -I \end{bmatrix} \hat{H}_N^{-1} \hat{g}_N. \tag{6.2.11}$$

设在解 x^* 处的 Lagrange 乘子为 λ^*，则有

$$g + Hx^* = A\lambda^*. \tag{6.2.12}$$

从而可知

$$\lambda^* = A_B^{-1}[g_B + H_{BB} x_B^* + H_{BN} x_N^*]. \tag{6.2.13}$$

如果 $\hat{H}_N$ 半正定，则在

$$(I - \hat{H}_N \hat{H}_N^+) \hat{g}_N = 0 \tag{6.2.14}$$

时，问题(6.2.5)有界，且(6.2.5)的解可表示为

$$x_N^* = -\hat{H}_N^+ \hat{g}_N + (I - \hat{H}_N^+ \hat{H}_N)\tilde{x}. \qquad (6.2.15)$$

$\tilde{x} \in \mathbf{R}^m$ 是任何向量，H^+ 表示 H 的广义逆矩阵. 于是利用(6.2.15)和(6.2.4)可求得原问题(6.2.1)与(6.2.2)的解. 如果(6.2.14)不成立，显然可见问题(6.2.5)无下界，从而原问题也无下界. 如果 $\hat{H}_N$ 有负特征值，则(6.2.5)没有下界. 故知此时问题(6.2.1)与(6.2.2)也无有限解.

消去法简单、直观，它的不足之处是：A_B 可能接近一奇异阵，此时由(6.2.11)求解 x^* 将会使数值不稳定.

消去法的一个直接推广是广义消去 (generalized elimination)法. 设 $y_1, \cdots, y_m$ 是 Range(A)中的一组线性无关向量；$z_1, \cdots, z_{n-m}$ 是 Null(A^T) 中的一组线性无关向量，记 $Y = [y_1, \cdots, y_m]$，$Z = [z_1, \cdots, z_{n-m}]$，则有 A^TY 非奇异，$A^TZ = 0$. 于是，(6.2.2)的解可表为

$$x = Y(A^TY)^{-1}b + Z\hat{x}. \qquad (6.2.16)$$

其中 $\hat{x} \in \mathbf{R}^{n-m}$ 是自由变量. 将(6.2.16)代入(6.2.1)式，就得到

$$\min (g + HY(A^TY)^{-1}b)^T Z\hat{x} + \frac{1}{2}\hat{x}^T Z^T H Z\hat{x}. \qquad (6.2.17)$$

假定 Z^THZ 正定，则由(6.2.17)可解出 $\hat{x}^*$：

$$\hat{x}^* = -(Z^THZ)^{-1}Z^T(g + HY(A^TY)^{-1}b). \qquad (6.2.18)$$

由(6.2.18)和(6.2.16)即可得到(6.2.1)与(6.2.2)的解：

$$x^* = Y(A^TY)^{-1}b - Z(Z^THZ)^{-1}Z^T(g + HY(A^TY)^{-1}b). \qquad (6.2.19)$$

于是，相应的 Lagrange 乘子可表示为

$$\begin{aligned} \lambda^* &= (A^TY)^{-T}Y^T[g + Hx^*] \\ &= (A^TY)^{-T}Y^T[I - HZ(Z^THZ)^{-1}Z^T]g \\ &\quad + (A^TY)^{-1}Y^TH[I - Z(Z^THZ)^{-1}Z^TH]Y(A^TY)^{-1}b. \end{aligned} \qquad (6.2.20)$$

如果适当选取 Y，我们可使

$$A^T Y = I, \tag{6.2.21}$$

此时，(6.2.19)与(6.2.20)变成

$$x^* = Yb - Z(Z^T HZ)^{-1} Z^T [g + HYb], \tag{6.2.22}$$

$$\lambda^* = Y^T [Pg + HP^T Yb]. \tag{6.2.23}$$

其中 P 是一个从 $\mathbf{R}^n$ 到 Range(A)的仿射映照:

$$P = I - HZ(Z^T HZ)^{-1} Z^T. \tag{6.2.24}$$

显然(6.2.2)的可行域是一个与 Null(A^T)平行的子空间. 广义消去法正是基于在这子空间上找二次函数的极小. 利用 Z 中的列向量 $z_i(i=1, \cdots, n-m)$作为基向量，我们将二次函数 $Q(x)$ 在子空间上求极小写成在子空间上的一个无约束二次函数极小问题(6.2.17). 从而称 $Z^T HZ$ 为既约(reduced)海色阵，称向量 $Z^T(g + HY(A^T Y)^{-1} b)$为既约梯度.

取

$$Y = \begin{bmatrix} A_B^{-1} \\ 0 \end{bmatrix}, \tag{6.2.25}$$

$$Z = \begin{bmatrix} -A_B^{-T} A_N^T \\ I \end{bmatrix}. \tag{6.2.26}$$

则知 (6.2.21) 成立，这时由(6.2.22)与(6.2.23)就可得到(6.2.11)和(6.2.13).

另一种选取 Y 和 Z 的方法是基于 A 的 QR 分解. 设

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix} = [Q_1 \quad Q_2]\begin{bmatrix} R \\ 0 \end{bmatrix}. \tag{6.2.27}$$

其中 Q 是正交阵，$R \in \mathbf{R}^{m\times m}$ 是非奇异上三角阵. 于是我们可取

$$Y = (A^+)^T = Q_1 R^{-T}, \tag{6.2.28}$$

$$Z = Q_2. \tag{6.2.29}$$

显然(6.2.21)也成立.

$A^T Z = 0$ 和(6.2.21)可写成如下形式

$$A^T [Y \quad Z] = [I \quad 0]. \tag{6.2.30}$$

所以对任何 $V \in \mathbf{R}^{n\times(n-m)}$ 使得$[A \quad V]$为非奇异，不难看出

$$[Y \quad Z]=\begin{bmatrix} A^T \\ V^T \end{bmatrix}^{-1} \tag{6.2.31}$$

必满足(6.2.31).

在(6.2.32)中取$V=\begin{bmatrix} 0 \\ I_N \end{bmatrix}$，则得到(6.2.25)与(6.2.26)；取$V=Q_2$，则得到(6.2.28)与(6.2.29). V 的选取方式不同，则得到不同的 Y 和 Z，从而导致不同的计算 x^* 和 λ^* 的公式. 正由于这个一般形式(6.2.31)，Fletcher(1987)指出了不少早期求解二次规划的方法，如 Beale (1959)方法、Wolfe (1963) 方法和 Murray (1971)方法，都可以看成是广义消去法.

我们可将(6.2.22)与(6.2.23)写成

$$x^*=-Z(Z^THZ)^{-1}Zg+P^TYb, \tag{6.2.32}$$

$$\lambda^*=Y^TPg+Y^THP^TYb. \tag{6.2.33}$$

下面给出的 Lagrange 方法也可将解表示成(6.2.32)与(6.2.33)的形式. Lagrange 方法是求解可行域内的 Kuhn-Tucker 点，即 Lagrange 函数的稳定点. 对于问题(6.2.1)与(6.2.2)，我们要求

$$g+Hx=A\lambda, \tag{6.2.34}$$

$$A^Tx=b. \tag{6.2.35}$$

写成矩阵形式即为

$$\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix}=-\begin{bmatrix} g \\ b \end{bmatrix}. \tag{6.2.36}$$

设矩阵

$$\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix} \tag{6.2.37}$$

可逆，则存在矩阵 $U\in\mathbf{R}^{n\times n}$，$W\in\mathbf{R}^{n\times m}$，$T\in\mathbf{R}^{m\times m}$，使得

$$\begin{bmatrix} U & W \\ W^T & T \end{bmatrix}=\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix}^{-1}. \tag{6.2.38}$$

于是从(6.2.36)和(6.2.38)可知 Lagrange 函数的稳定点(x^*, λ^*)为

$$\begin{bmatrix} x^* \\ \lambda^* \end{bmatrix} = -\begin{bmatrix} U & W \\ W^T & T \end{bmatrix} \begin{bmatrix} g \\ b \end{bmatrix}. \tag{6.2.39}$$

从而可知

$$x^* = -Ug - Wb, \tag{6.2.40}$$

$$\lambda^* = -W^T g - Tb. \tag{6.2.41}$$

只要矩阵(6.2.37)可逆，则(6.2.38)中的 U、W、T 唯一确定. 但 U、W、T 的表达形式有不少方式，从而导出不同形式的计算公式(6.2.40)与(6.2.41).

当 H 可逆，A 列满秩，则 $(A^T H^{-1} A)^{-1}$ 存在，且有

$$\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H^{-1} - H^{-1}A(A^T H A)^{-1} A^T H^{-1} & -H^{-1}A(A^T H^{-1} A)^{-1} \\ -(A^T H^{-1} A)^{-1} A^T H^{-1} & -(A^T H^{-1} A)^{-1} \end{bmatrix}. \tag{6.2.42}$$

利用(6.2.42)和(6.2.39)即得到

$$x^* = -H^{-1}g + H^{-1}A(A^T H^{-1} A)^{-1}[A^T H^{-1} g + b], \tag{6.2.43}$$

$$\lambda^* = (A^T H^{-1} A)^{-1}[A^T H^{-1} g + b]. \tag{6.2.44}$$

设 Y、Z 由(6.2.31)定义，则不难证明

$$\begin{bmatrix} H & -A \\ -A^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} Z(Z^T H Z)^{-1} Z^T & -P^T Y \\ -Y^T P & -Y^T H P^T Y \end{bmatrix}. \tag{6.2.45}$$

于是，解(6.2.40)、(6.2.41)可写成(6.2.32)、(6.2.33). 从而可看出 Lagrange 方法和广义消去法的等价性.

利用 Lagrange 方法求解二次规划问题的关键在于如何有效地求解线性方程组(6.2.36)，有时等价于求解矩阵(6.2.37)的逆. 关于这方面的详细讨论可见 Fletcher(1987).

§ 6.3 积 极 集 法

积极集法是通过求解有限个等式约束二次规划问题来解决一

般约束下的二次规划问题. 方法的理论基础是下面的引理:

引理 **6.3.1** 设 x^* 是二次规划问题(6.0.1)～(6.0.3)的局部解,则 x^* 也必是问题

$$\min_{x\in\mathbb{R}^n} g^Tx+\frac{1}{2}x^THx \tag{6.3.1}$$

$$\text{s.t.}\quad a_i^Tx=b_i,\ i\in I(x^*)\cup E. \tag{6.3.2}$$

的解. 反之,如果 x^* 是(6.0.1)～(6.0.3)的可行点,且是(6.3.1)与(6.3.2)的局部解,如果在 x^* 处的 Lagrange 乘子 λ^* 满足

$$\lambda_i^*\geqslant 0,\ i\in I(x^*). \tag{6.3.3}$$

则 x^* 必是原问题(6.0.1)～(6.0.3)的 Kuhn-Tucker 点.

证明 由于在 x^* 附近处,(6.3.2)的可行点必是(6.0.2)与(6.0.3)的可行点,所以显然当 x^* 是问题(6.0.1)～(6.0.3)的局部解时,x^* 也必是问题(6.3.1)与(6.3.2)的局部解.

现在假定 x^* 是(6.3.1)与(6.3.2)的局部解且满足(6.0.2)与(6.0.3),如果(6.3.3)满足,我们定义

$$\lambda_i^*=0,\ i\in\{m_e+1,\ \cdots,\ m\}\setminus I(x^*). \tag{6.3.4}$$

则知 x^* 是原问题(6.0.1)～(6.0.3)的 Kuhn-Tucker 点. ▌

积极集法是一个可行点方法,即每个迭代点都要求是可行点. 积极集法的基本思想是每次迭代求解一个等式约束的二次规划问题. 如果(6.3.3)得到满足,则停止计算. 否则,可去掉一个约束,重新求解等式约束问题,直到求解到(6.3.3)得到满足.

在第 k 次迭代, 有可行点 x_k 以及一个下标集合 $S_k\subset E\cup I$. 其中 $E=\{1,\ \cdots,\ m_e\}$, $I=\{m_e+1,\ \cdots,\ m\}$. 设 d_k 是问题

$$\min_{d\in\mathbb{R}^n} g^T(x_k+d)+\frac{1}{2}(x_k+d)^TH(x_k+d), \tag{6.3.5}$$

$$\text{s.t.}\quad a_i^Td=0,\quad i\in S_k \tag{6.3.6}$$

的解. $\lambda_i^{(k)}(i\in S_k)$是相应的 Lagrange 乘子. 如果 $d_k=0$, 则 x_k 是问题

$$\min_{x\in\mathbb{R}^n} g^Tx+\frac{1}{2}x^THx, \tag{6.3.7}$$

$$a_i^Tx=b_i,\ i\in S_k \tag{6.3.8}$$

的 Kuhn-Tucker 点. 如果 $\lambda_i^{(k)}\geqslant 0$ 对一切 $i\in S_k\cap I$ 均成立, 则 x_k 是原问题(6.0.1)～(6.0.3)的 Kuhn-Tucker 点. 否则, 我们可令 $i_k\in S_k\cap I$ 使得

$$\lambda_{i_k}^{(k)}=\min_{i\in S_k\cap I}\lambda_i^{(k)}<0. \tag{6.3.9}$$

且取 $S_k:=S_k\setminus\{i_k\}$, 然后重新求解(6.3.5)与(6.3.6).

设(6.3.5)与(6.3.6)的解 $d_k\neq 0$. 这时, x_k+d_k 很可能不是原问题(6.0.1)～(6.0.3)的可行点. 我们取线段 $x_kx_k+d_k$ 上靠 x_k+d_k 最近的可行点作为下次迭代的迭代点 x_{k+1}. 也就是说:

$$x_{k+1}=x_k+\alpha_kd_k, \tag{6.3.10}$$

其中

$$\alpha_k=\min\left\{1,\ \min_{\substack{i\notin S_k\\ a_i^Td_k<0}}\frac{b_i-a_i^Tx_k}{a_i^Td_k}\right\}. \tag{6.3.11}$$

现在我们可给出积极集法的主要步骤, 这个方法是由 Fletcher(1971)提出来的.

算法 6.3.2

步 1　给出可行点 x_1, 令 $S_1=E\cup I(x_1)$; $k:=1$;

步 2　求解(6.3.5)与(6.3.6), 得出 d_k.

　　如果 $d_k\neq 0$, 则转步 3;

　　如果 $\lambda_i^{(k)}\geqslant 0(i\in S_k\cap I)$, 则停;

　　由(6.3.7)求得 i_k;

　　$S_k:=S_k\setminus\{i_k\}$, 转步 2;

步 3　由(6.3.11)计算步长 α_k.

$$x_{k+1}:=x_k+\alpha_kd_k,$$

　　对 $j\notin S_k$ 作:

　　如果 $a_j^Tx_{k+1}=b_j$, 则 $S_k:=S_k\cup\{j\}$;

步 4　$S_{k+1}:=S_k$; $k:=k+1$, 转步 2.

显然算法 6.3.2 是一个下降算法, 即有

$$Q(x_{k+1})\leqslant Q(x_k). \tag{6.3.12}$$

且只要 $d_k\neq 0$(x_k 不是(6.3.7)与(6.3.8) 的 Kuhn-Tucker 点) 且 $\alpha_k>0$, 则有

$$Q(x_{k+1})<Q(x_k). \tag{6.3.13}$$

假定算法(6.3.2)不有限终止，则由于只有有限个约束，不可能无穷次的增加 S_k 的元素个数而不减少．故必有无穷多个 k 使得 $d_k=0$．于是有无穷多个 k 使得 x_k 是(6.3.7)与(6.3.8)的解．由于只有有限个约束，S_k 只可能有有限个不同的集合．从而可知必有无限个 x_k 相等，因为算法是下降算法，不难证明，存在充分大的 k_0，当 $k\geqslant k_0$ 时，

$$Q(x_{k+1})=Q(x_k). \tag{6.3.14}$$

故知 $\alpha_k=0$ 对一切 $k\geqslant k_0$ 均成立，即有

$$x_k=x_{k_0}=\bar{x} \quad \forall k\geqslant k_0. \tag{6.3.15}$$

于是算法将无穷循环．设 $k_1>k_0$ 使得

$$d_{k_1}=0, \tag{6.3.16}$$

则存在 $\lambda_i(i\in S_{k_1})$，使得

$$g+H\bar{x}=\sum_{i\in S_{k_1}} a_i\lambda_i^{(k)}. \tag{6.3.17}$$

不妨设在去掉 i_{k_1} 后产生的 $\bar{d}_{k_1}$ 不等于零，则由于 $\alpha_{k_1}=0$ 知存在

$$j\notin S_{k_1}/\{i_{k_1}\}, \tag{6.3.18}$$

使得

$$a_j^T x=b_j, \tag{6.3.19}$$

$$a_j^T\bar{d}_{k_1}<0. \tag{6.3.20}$$

其中 $\bar{d}_{k_1}$ 是问题

$$\min g^T(d+\bar{x})+\frac{1}{2}(d+\bar{x})^T H(d+\bar{x}), \tag{6.3.21}$$

$$\text{s.t.} \quad a_i^T d=0,\ i\in S_{k_1}\setminus\{i_{k_1}\} \tag{6.3.22}$$

的解．显然

$$(g+H\bar{x})^T\bar{d}_{k_1}<0. \tag{6.3.23}$$

利用(6.3.17)、(6.3.22)和(6.3.23)可知

$$\lambda_{i_{k_1}}^{(k_1)}a_{i_{k_1}}^T\bar{d}_{k_1}<0. \tag{6.3.24}$$

由 $\lambda_i^{(k)}$ 的定义知 $\lambda_i^{(k)}<0$，故有

$$a_{i_{k_1}}^T\bar{d}_{k_1}>0. \tag{6.3.25}$$

从(6.3.20)和(6.3.25)知 $j\neq i_{k_1}$，再利用关系式(6.3.15)即知

$$j \notin S_{k_1}. \tag{6.3.26}$$

设 $k_2 > k_1$ 是使得 $d_k = 0$ 的最小整数，则知

$$j \in S_{k_2}. \tag{6.3.27}$$

所以，我们有

$$g + H\bar{x} = \sum_{i \in S_{k_2}} a_i \lambda_i^{(k_2)}, \tag{6.3.28}$$

且 $S_{k_2} \neq S_{k_1}$. 从而 $\bar{x}$ 是两个不同的等式约束优化的 Kuhn-Tucker 点. 对这种情况，我们称为退化(degeneracy)，它和线性规划的退化情形是相似的. 关于线性规划的退化情形的讨论，可见 Fletcher (1987)或者 Luenberger(1984). 克服退化的一种方法是对约束进行小的扰动，这种利用扰动的技巧是由 Charnes(1952)提出来的；另一种方法是由 Dantzig、Orden 和Wolfe(1955)提出的字典顺序 (lexicographic order)法，这种方法避免相同的 S_k 重新出现. 关于处理退化的技巧的详细讨论可见 Fletcher(1987).

从上面的讨论，我们有以下定理：

定理 6.3.3 设点列$\{x_k\}$由算法 6.3.2 产生. 若算法不是有限终止，则必出现退化情形. 若算法 6.3.2 加上克服退化的技巧，则算法必有限次迭代后终止于问题(6.0.1)～(6.0.3)的 Kuhn-Tucker 点.

关于算法 6.3.2 所需要的初始可行点，可利用求解线性规划的单纯形(simplex)方法的第一阶段(Phase I)方法求得.

另一种可能是问题(6.3.7)与(6.3.8)无下界，此时，我们可求得一方向 d_k，使得

$$a_i^T d_k = 0, \ i \in S_k, \tag{6.3.29}$$

且有

$$d_k^T H d_k < 0, \tag{6.3.30}$$

或者

$$(g^T + Hx_k)^T d_k < 0, \ d_k^T H d_k = 0. \tag{6.3.31}$$

如果对一切 $i \notin S_k$ 均有 $a_i^T d_k \geq 0$，则不难看出原问题(6.0.1)～(6.0.3)无下界. 否则，必有 $i \notin S_k$ 且 $a_i^T d_k < 0$. 从而可知

$$x_k + \alpha d_k$$

对充分大的 α 不满足(6.0.3). 于是，我们可取 α_k 尽可能大且

$x_k+\alpha_k d_k$ 是可行点.

§6.4 对偶方法

本节我们讨论严格凸的二次规划问题:

$$\min_{x\in\mathbb{R}^n} g^T x+\frac{1}{2}x^T H x=Q(x), \tag{6.4.1}$$

$$\text{s.t.}\quad a_i^T x=b_i,\ i\in E, \tag{6.4.2}$$

$$a_i^T x\geqslant b_i,\ i\in I. \tag{6.4.3}$$

的对偶方法. H 是对称正定阵, 故 $Q(x)$ 是严格凸函数. 考虑(6.4.1)～(6.4.3)的对偶问题(6.1.27)与(6.1.28)的积极集法. 设有一个可行点 $\lambda^{(k)}$ 是

$$\max (b+AH^{-1}g)^T\lambda-\frac{1}{2}\lambda^T(A^TH^{-1}A)\lambda, \tag{6.4.4}$$

$$\text{s.t.}\quad \lambda_i=0,\ i\in \bar{S}_k \tag{6.4.5}$$

的 Kuhn-Tucker 点, 则令

$$x_k=-H^{-1}[g-A\lambda^{(k)}], \tag{6.4.6}$$

于是知 x_k 是问题

$$\min g^T x+\frac{1}{2}x^T H x, \tag{6.4.7}$$

$$\text{s.t.}\quad a_i x=b_i,\ i\notin \bar{S}_k \tag{6.4.8}$$

的 Kuhn-Tucker 点. 由于求解(6.1.27)与(6.1.28)的积极集法正是逐步求解(6.4.4)与(6.4.5), 解(6.4.1)～(6.4.3)的对偶算法正是等价地逐步求解(6.4.7)与(6.4.8), 且 $\bar{S}_k$ 的确定是由积极集法应用于对偶问题而确定的.

记 $S_k=I\cup E\backslash \bar{S}_k$, 则(6.4.7)、(6.4.8)和(6.3.7)、(6.3.8)是一样的. 我们记列向量由 $a_i(i\in S_k)$ 组成的矩阵为 A_k, 且

$$A_k^*=(A_k^TH^{-1}A_k)^{-1}A_k^TH^{-1}, \tag{6.4.9}$$

$$\hat{H}_k=H^{-1}(I-A_kA_k^*). \tag{6.4.10}$$

利用这些记号, 我们可将 Goldfarb 和 Idnani(1983)的对偶方法叙述如下($m_e=0$ 的情形):

算法 6.4.1

步 1 $$x_1=-H^{-1}g,\ f_1=\frac{1}{2}g^Tx_1;$$
$$S_1=\Phi,\qquad k:=1;$$
$$\lambda_1=\Phi,\qquad q=0.$$

步 2 计算 $r_i=b_i-a_i^Tx_k,\ i=1,\ \cdots,\ n$;

如果 $r_i\leqslant 0$, 则停;

令 p 使得 $r_p>0$; $\lambda_k:=\begin{pmatrix}\lambda_k\\0\end{pmatrix}$;

步 3 $$d_k:=\hat{H}_ka_p;$$
$$y_k:=A_k^*a_p.$$

令

$$\alpha_k=\min_{\substack{(y_k)_j>0\\(\lambda_k)_j\neq 0}}\left\{\frac{(\lambda_k)_j}{(y_k)_j}\right\}=\frac{(\lambda_k)_l}{(y_k)_l}. \tag{6.4.11}$$

步 4 如果 $d_k\neq 0$ 则转步 5.

如果 $\alpha_k=\infty$, 则停(原问题无可行点);

$$S_k:=S_k\setminus\{l\};\ q=q-1;$$
$$\lambda_k:=\lambda_k+\alpha_k\begin{pmatrix}-y_k\\1\end{pmatrix}.$$

修改 A_k^* 和 $\hat{H}_k$; 转步 3.

步 5 $\hat{\alpha}:=-(b_p-a_p^Tx_k)/a_p^Td_k$;

$$\alpha_k:=\min\{\alpha_k,\ \hat{\alpha}\},$$
$$x_{k+1}=x_k+\alpha_kd_k,$$
$$f_{k+1}=f_k+\alpha_ka_p^Td_k\left(\frac{1}{2}\alpha_k+(\lambda_k)_{q+1}\right),$$
$$\lambda_{k+1}:=\lambda_k+\alpha_k\begin{pmatrix}-y_k\\1\end{pmatrix}.$$

步 6 如果 $\alpha_k<\hat{\alpha}$, 则转步 7.

$$S_{k+1}=S_k\cup\{p\};\ q=q+1;$$

修改 $\hat{H}_{k+1}$ 和 A_{k+1}^*; $k=k+1$; 转步 2.

步 7 $S_{k+1}:=S_k\setminus\{l\};\ q=q-1;$

从 λ_k 中去掉第 l 个分量，得到 λ_{k+1}；

修改 $\hat{H}_k$ 和 A_k^*；转步 3.

由于 λ_k 是对偶问题的可行点，故必有

$$(\lambda_k)_i \geqslant 0, \ i \in I. \tag{6.4.12}$$

于是在对偶算法6.4.1中产生的 x_k 不仅是问题(6.3.7)与(6.3.8)的解，而且也是

$$\min_{x \in \mathbb{R}^n} g^T x + \frac{1}{2} x^T H x, \tag{6.4.13}$$

$$\text{s.t.} \quad a_i x = b_i, \ i \in S_k \cap E, \tag{6.4.14}$$

$$a_i^T x \geqslant b_i, \ i \in S_k \cap I. \tag{6.4.15}$$

的解.

在具体计算过程中，Goldfarb 和 Idnani 建议用 H 的 Cholesky 分解

$$H = LL^T \tag{6.4.16}$$

以及对矩阵 $L^{-1}A_k$ 进行 QR 分解，即有

$$L^{-1}A_k = Q_k \begin{bmatrix} R_k \\ 0 \end{bmatrix}. \tag{6.4.17}$$

这样做比直接利用 H^{-1} 数值稳定性要好多了.

Powell(1985)发现Goldfarb和Idnani的分解方式(6.4.16)与(6.4.17)仍可能出现数值不稳定，于是他建议采用

$$A_k = Q_k \begin{bmatrix} R_k \\ 0 \end{bmatrix} = [Q_k^{(1)} \quad Q_k^{(2)}] \begin{bmatrix} R_k \\ 0 \end{bmatrix}. \tag{6.4.18}$$

然后考虑既约 (reduced) Hessian 阵 $[Q_k^{(2)}]^T H Q_k^{(2)}$ 的反 Cholesky 分解，即

$$U_k U_k^T = [Q_k^{(2)}]^T H Q_k^{(2)}.$$

其中 U_k 是上三角矩阵. Powell 给出的算法每次迭代修正 $Q_k^{(1)}$, R_k 和 U_k.

另一个对偶算法是由 Lemke(1962)给出的. Lemke 方法的基本思想是利用 Beale (1959)的共轭方向法来求解对偶问题(6.1.27)与(6.1.28).

对偶算法实质上就是对对偶问题用一个原始(prime)方法来求解．对偶算法的一个优点是它不需要利用第一阶段(phase I)方法去求可行点．对偶问题的一个显而易见的可行点是$\lambda=0$．

§6.5　线性互补问题

本节介绍一种解决二次规划问题

$$\min_{x\in\mathbb{R}^n} g^Tx+\frac{1}{2}x^THx, \tag{6.5.1}$$

$$\text{s.t.}\quad A^Tx\geqslant b, \tag{6.5.2}$$

$$x\geqslant 0 \tag{6.5.3}$$

的特殊方法．这种方法是将二次规划(6.5.1)～(6.5.3)化成一个线性互补问题，然后对线性互补问题求解．

不难求得(6.5.1)～(6.5.3)的 Kuhn-Tucker 条件为

$$g+Hx=A\lambda+\mu, \tag{6.5.4}$$

$$A^Tx=b+r, \tag{6.5.5}$$

$$\lambda\geqslant 0,\ \mu\geqslant 0,\ x\geqslant 0,\ r\geqslant 0. \tag{6.5.6}$$

我们定义

$$M=\begin{bmatrix} H & -A \\ A^T & 0 \end{bmatrix},\quad q=\begin{bmatrix} g \\ -b \end{bmatrix}. \tag{6.5.7}$$

则由(6.5.4)～(6.5.6)得到

$$w-Mz=q, \tag{6.5.8}$$

$$w\geqslant 0,\ z\geqslant 0,\ w^Tz=0. \tag{6.5.9}$$

其中

$$w=\begin{bmatrix} \mu \\ r \end{bmatrix},\quad z=\begin{bmatrix} x \\ \lambda \end{bmatrix}. \tag{6.5.10}$$

问题(6.5.8)与(6.5.9)被称为线性互补问题 (Linear Complementary probem)，简称为 LCP.

不难看出，当$q\geqslant 0$时，线性互补问题的解为$w=q$, $z=0$. 求解线性互补问题的方法都是试图找到一个$\{1, \cdots, n+m\}$的子集S_w，然后通过线性变换，使得(6.5.8)变成形式：

$$w_k - M_k z_k = q_k, \tag{6.5.11}$$

$$w_k \geqslant 0,\ z_k \geqslant 0,\ w_k^T z_k = 0. \tag{6.5.12}$$

其中

$$(w_k)_i = \begin{cases} w_i, & \text{如果 } i \in S_k; \\ z_i, & \text{否则.} \end{cases} \tag{6.5.13}$$

$$(z_k)_i = \begin{cases} z_i, & \text{如果 } i \in S_k; \\ w_i, & \text{如果 } i \notin S_k. \end{cases} \tag{6.5.14}$$

集合 S_k 的选择是设法使得 $\min\limits_{1\leqslant j\leqslant n+m} (q_k)_j$ 增加. 每次迭代 w_k 和 z_k 互相交换某一相应位置的分量. 这实质上是把 w_k 看成是基变量, 而把 z_k 看成非基变量, 所以方法和线性规划的单纯形法有着十分相似之处.

解 LCP 问题的经典方法有主元消去法(principal pivoting method), 也称为 Dantzig-Wolfe 方法. 不难证明, Dantzig-Wolfe 方法实质上等价于我们在 6.3 节介绍的积极集法. 关于 Dantzig-Wolfe 方法的详细讨论可见 Fletcher (1987). 另一个有名的求解 LCP 问题的方法是 Lemke (1965)方法, 它是基于求解

$$w - Mz - z_0 e = q, \tag{6.5.15}$$

$$w \geqslant 0,\ z \geqslant 0,\ z_0 \geqslant 0,\ w^T z = 0. \tag{6.5.16}$$

其中 $e = (1, 1, \cdots, 1)^T$. Lemke 方法是逐步迭代使 $z_0 = 0$ 且 $q \geqslant 0$, 从而得到原来 LCP 问题的解. 在(6.5.15)中引入 $z_0 e$ 的好处在于很容易得到(6.5.15)与(6.5.16)的一个可行点

$$z_0 = -\min_{1\leqslant i\leqslant n+m} (q)_i, \tag{6.5.17}$$

$$z = 0,$$

$$w = q + z_0 e. \tag{6.5.18}$$

关于 Lemke 方法的介绍可见 Bazara 和 Shetty(1979).

线性互补问题是一个重要的问题, 它除了可用来求解二次规划的问题以外, 还可应用于对策论(Game theory)以及边值问题计算等问题. 它还与变分不等式有着密切的联系. 关于线性互补

问题的最新研究可参阅 Cottle, Kyparisis 和 Pang(1990).

§6.6 内点算法

Karmarkar 在 1984 年提出一个求解线性规划问题的内点算法. Karmarkar 算法不仅具有较好的理论性质,即它是多项式时间算法(polynomial time algorithm),而且大量的数值例子已表明它是一个十分有效的实用方法. 所以不少人开始对内点法作进一步研究,并将 Karmarkar 算法的基本思想推广到非线性规划.

考虑凸规划问题

$$\min_{x\in\mathbf{R}^n} g^T x+\frac{1}{2}x^T Hx=Q(x), \tag{6.6.1}$$

$$\text{s.t.}\quad Ax=b, \tag{6.6.2}$$

$$x\geqslant 0, \tag{6.6.3}$$

设 x_k 是一个内点,即

$$Ax_k=b, \tag{6.6.4}$$

$$x_k>0. \tag{6.6.5}$$

与 Karmarkar 方法一样,定义矩阵

$$D_k=\begin{bmatrix}(x_k)_1 & & 0\\ & \ddots & \\ 0 & & (x_k)_n\end{bmatrix}. \tag{6.6.6}$$

作变量代换 $\hat{x}:=T_k x$

$$\hat{x}_i=\frac{(n+1)D_k^{-1}x}{e^T D_k^{-1}x+1},\quad i=1,\cdots,n; \tag{6.6.7}$$

$$\hat{x}_{n+1}=(n+1)/[e^T D_k^{-1}x+1]. \tag{6.6.8}$$

则将问题(6.6.1)~(6.6.3)化成

$$\min_{\hat{x}\in\mathbf{R}^{n+1}} \hat{x}_{n+1}Q(T^{-1}\hat{x}), \tag{6.6.9}$$

$$\text{s.t.}\ AD_k\hat{x}[n]-\hat{x}_{n+1}b=0, \tag{6.6.10}$$

$$e^T\hat{x}=n+1, \tag{6.6.11}$$

$$\hat{x}[n]\geqslant 0,\ \hat{x}_{n+1}>0. \tag{6.6.12}$$

其中 $e=(1,\ \cdots,\ 1)^T$; $\hat{x}[n]=(\hat{x}_1,\ \cdots,\ \hat{x}_n)^T$.

正是基于(6.6.9)～(6.6.12), Ye 和 Tse(1989)提出的一个解二次规划问题(6.6.1)～(6.6.3)的方法,在每次迭代时求解

$$\min\ \hat{g}_k^T\hat{x}[n]+\frac{1}{2}\hat{x}[n]^T\hat{H}_k\hat{x}[n]/\hat{x}_{n+1}, \tag{6.6.13}$$

$$\text{s.t.}\quad \hat{A}_k\hat{x}=\hat{b}, \tag{6.6.14}$$

$$\|\hat{x}-e\|_2\leqslant\beta<1. \tag{6.6.15}$$

其中

$$\hat{H}=D_kHD_k, \tag{6.6.16}$$

$$\hat{g}=D_kg, \tag{6.6.17}$$

$$\hat{A}_k=\begin{bmatrix} AD_k & -b \\ & e^T \end{bmatrix}, \tag{6.6.18}$$

$$\hat{b}=\begin{bmatrix} 0 \\ \vdots \\ 0 \\ n+1 \end{bmatrix}. \tag{6.6.19}$$

$\beta<1$ 是一个与 k 无关的正常数. 利用 Kuhn-Tucker 定理, 我们知求解(6.6.13)与(6.6.14)等价于

$$\hat{g}_k+\hat{x}_{n+1}^{-1}\hat{H}_k\hat{x}[n]=\hat{A}_k[n]^T\lambda+\mu(\hat{x}[n]-e) \tag{6.6.20}$$

$$-\frac{1}{2\hat{x}_{n+1}^2}\hat{x}[n]^T\hat{H}_k\hat{x}[n]=(\hat{a}_{n+1}^{(k)})^T\lambda+\mu(\hat{x}_{n+1}-1)=0, \tag{6.6.21}$$

$$\hat{A}_k\hat{x}=\hat{b}, \tag{6.6.22}$$

$$\|\hat{x}-e\|_2\leqslant\beta, \tag{6.6.23}$$

$$\mu[\|\hat{x}-e\|_2-\beta]=0,\ \mu\leqslant 0. \tag{6.6.24}$$

其中 $\hat{A}_k[n]$ 是 $\hat{A}_k$ 的前 n 列组成的矩阵, $a_{n+1}^{(k)}$ 是 $\hat{A}_k$ 的第 $n+1$ 列. 我们可将(6.6.20)和(6.6.22)写成矩阵形式

$$P_k\begin{bmatrix}\hat{x}[n]\\ \hat{\lambda}\end{bmatrix}=\hat{x}_{n+1}\bar{b}+\tilde{b}, \tag{6.6.25}$$

其中

$$P_k=\begin{bmatrix}\hat{H}_k+\hat{\mu}I & -\hat{A}[n]^T\\ \hat{A}[n] & 0\end{bmatrix}, \tag{6.6.26}$$

$$\bar{b}=\begin{bmatrix}-\hat{g}\\ b\\ -1\end{bmatrix}\quad \tilde{b}=\begin{bmatrix}\hat{\mu}e\\ 0\\ n+1\end{bmatrix}, \tag{6.6.27}$$

且

$$\hat{\lambda}=\hat{x}_{n+1}\lambda,\ \hat{\mu}=-\hat{x}_{n+1}\mu. \tag{6.6.28}$$

于是,对任何给定的 $\hat{\mu}\geqslant 0$. 我们可由(6.6.25)求得 $\hat{\lambda}$ 和 $\hat{x}[n]$. 将求出的 $\hat{\lambda}$ 和 $\hat{x}[n]$代入(6.6.21)式即可求得 $\hat{x}_{n+1}$. 于是我们对任何 $\hat{\mu}>0$, 可求得 $\hat{x}(\hat{\mu})$, 定义函数

$$h(\hat{\mu})=\|\hat{x}(\hat{\mu})-e\|-\beta. \tag{6.6.29}$$

Ye 和 Tse(1989)建议用对分法求得 $\hat{\mu}_k$ 使得 $h(\hat{\mu}_k)=0$ 成立,从而得到(6.6.13)～(6.6.15)的解 $\hat{x}(\hat{u}_k)$. 将 $\hat{x}(\hat{\mu}_k)$变换回去,即得到下一次迭代点 x_{k+1}, 也就是说

$$x_{k+1}=T_k^{-1}\hat{x}(\hat{u}_k)=\frac{D_k\hat{x}(\hat{u}_k)[n]}{\hat{x}(\hat{u}_k)_{n+1}}, \tag{6.6.30}$$

其中 $\hat{x}(\hat{\mu}_k)[n]=(\hat{x}(\hat{\mu}_k)_1,\cdots,\hat{x}(\hat{\mu}_k)_n)^T$.

下面是 Ye 和 Tse(1989) 给出的求解二次规划 (6.6.1)～(6.6.3)的内点算法.

算法6.6.1

步 1　给出(6.6.2)与(6.6.3)的严格内点 x_1, $k:=1$.

步 2　求解(6.6.13)～(6.6.15)给出 $\hat{x}(u_k)$, 利用(6.6.30)计算 x_{k+1}.

步 3　如果 x_{k+1}是 $K-T$ 点则停;
　　$k:=k+1$; 转步 2.

另一个将 Karmarkar 方法推广到凸二次规划的内点算法是由 Kapoor 和 Vaidya (1986)提出的. 内点算法是一个依然十分活跃的研究方向,近来有不少工作, 如 Goldfarb 和 Liu(1991)

Monteiro 和 Adler(1989), Ye 和 Todd(1990), 以及 Mehrotra 和 Sun(1990)等等. 值得指出的是:内点算法除了 Karmarkar 算法类型的利用投影(projection)的方法外,近年来人们又提出了一类称为"中点轨迹追踪"(Center path-following)算法. 这类新的算法不需要作投影变换, 关于它的详细讨论可见 Sonnevend (1985)、Ye 和 Todd(1990).

第 7 章

罚 函 数 法

罚函数(penalty function)法是通过求解一个或多个罚函数的极小来求约束规划问题之解的方法. 它的基本思想是将约束问题非约束化. 早期的求解非线性规划问题(1.1.1)~(1.1.3)的方法都是罚函数法. 由于这些早期方法均是需要求解一串罚函数极小的方法, 故它们也被称为*序列无约束极小*(sequential unconstrained minimization)方法. 罚函数可分为外点罚函数、内点罚函数以及混合罚函数. 内点罚函数也称为*障碍*(barrier)*罚函数*. 如果罚函数的极小点和原问题(1.1.1)~(1.1.3)的解吻合时, 我们称该罚函数为*精确罚函数*(exact penalty function). 罚函数法简单实用, 它直接利用无约束优化的算法来求解约束优化问题. 但是罚函数法一般要求罚因子趋于无穷或者需要求解非光滑优化, 所以利用一般无约束优化算法时常有困难, 另外罚函数法的收敛速度通常也是很慢的. 由于 Karmarkar(1984)提出的一个求解线性规划的内点法实质上等价于一个罚函数法, 人们开始重新重视罚函数法. 解约束规划的其他方法在判别迭代点好坏时都用到某一罚函数, 所以对罚函数的研究对于求解约束规划问题有着重要的作用.

§7.1 早期罚函数

最早的罚函数是由 Courant(1943)提出的, 它只是对等式约束问题(1.1.1)与(1.1.2)提出的. Courant 罚函数可写成

$$P(x,\ \sigma)=f(x)+\sigma\sum_{i=1}^{m_e}[c_i(x)]^2, \qquad (7.1.1)$$

其中 $\sigma\geqslant 0$ 是一个罚因子. 罚函数 (7.1.1) 可推广到一般约束规划问题, 即

$$P(x,\ \sigma)=f(x)+\sigma\left[\sum_{i=1}^{m_e}[c_i(x)]^2+\sum_{i=m_e+1}^{m}[c_i(x)_-]^2\right], \quad (7.1.2)$$

其中

$$c_i(x)_-=\min\{0,\ c_i(x)\}. \quad (7.1.3)$$

我们定义 $\bar{c}(x)=(\bar{c}_1(x),\ \cdots,\ \bar{c}_m(x))^T$, 且

$$\bar{c}_i(x)=\begin{cases}c_i(x), & i\in E;\\ c_i(x)_-, & i\in I.\end{cases} \quad (7.1.4)$$

则(7.1.2)式可写成

$$P(x,\ \sigma)=f(x)+\sigma\|\bar{c}(x)\|_2^2. \quad (7.1.5)$$

显然, x 是问题(1.1.1)～(1.1.3)的可行点当且仅当

$$\bar{c}(x)=0, \quad (7.1.6)$$

所以 $\bar{c}(x)$ 被称为约束违反度(constraint violation). 不难看出

$$\|\bar{c}(x)\|_2=\operatorname{dist}(c(x),\ C), \quad (7.1.7)$$

其中 C 是可行集

$$C=\{c\,|\,c\in\mathbf{R}^m,\ c_i\geqslant 0,\ i\in I\}. \quad (7.1.8)$$

因为对一切 $x\in X$ (可行域), 均有 $c(x)\in C$, 所以 $\operatorname{dist}(c(x), C)$ 可看成是从点 x 到可行域 X 的一种度量.

我们记 $x(\sigma)$ 是问题

$$\min_{x\in\mathbf{R}^n} f(x)+\sigma\|\bar{c}(x)\|_2^2 \quad (7.1.9)$$

的解. 首先有如下引理:

引理 7.1.1　设 $\sigma_2\geqslant\sigma_1\geqslant 0$, 则必有

$$f(x(\sigma_2))\geqslant f(x(\sigma_1)), \quad (7.1.10)$$

$$\|\bar{c}(x(\sigma_2))\|_2\leqslant\|\bar{c}(x(\sigma_1))\|_2. \quad (7.1.11)$$

证明　由 $x(\sigma)$ 的定义, 我们有

$$f(x(\sigma_1))+\sigma_1\|\bar{c}(x(\sigma_1))\|_2^2 \leqslant f(x(\sigma_2))+\sigma_1\|\bar{c}(x(\sigma_2))\|_2^2, \quad (7.1.12)$$

$$f(x(\sigma_2))+\sigma_2\|\bar{c}(x(\sigma_2))\|_2^2 \leqslant f(x(\sigma_1))+\sigma_2\|\bar{c}(x(\sigma_1))\|_2^2, \quad (7.1.13)$$

从上两式即知

$$(\sigma_1-\sigma_2)[\|\bar{c}(x(\sigma_1))\|_2^2-\|\bar{c}(x(\sigma_2))\|_2^2]\leqslant 0, \tag{7.1.14}$$

于是(7.1.11)成立. 由(7.1.12)和(7.1.11)可推得

$$\begin{aligned}f(x(\sigma_1))&\leqslant f(x(\sigma_2))+\sigma_1[\|\bar{c}(x(\sigma_2))\|_2^2-\|\bar{c}(x(\sigma_1))\|_2^2]\\&\leqslant f(x(\sigma_2)),\end{aligned} \tag{7.1.15}$$

故知(7.1.10)成立. ▮

引理 7.1.2 给定 $\sigma\geqslant 0$, $x(\sigma)$ 是 (7.1.9)的解, 则 $x(\sigma)$也是约束问题

$$\min_{x\in\mathbf{R}^n} f(x), \tag{7.1.16}$$

$$\text{s.t.}\ \|\bar{c}(x)\|_2\leqslant\delta \tag{7.1.17}$$

的解, 其中 $\delta=\|\bar{c}(x(\sigma))\|_2$.

证明 对任何 x 满足(7.1.17), 由 $x(\sigma)$的定义, 我们有

$$f(x)+\sigma\|\bar{c}(x)\|_2^2\geqslant f(x(\sigma))+\sigma\|\bar{c}(x(\sigma))\|_2^2, \tag{7.1.18}$$

所以

$$\begin{aligned}f(x)&\geqslant f(x(\sigma))+\sigma[\|\bar{c}(x(\sigma))\|_2^2-\|\bar{c}(x)\|_2^2]\\&\geqslant f(x(\sigma)).\end{aligned} \tag{7.1.19}$$

故知 $x(\sigma)$是问题(7.1.16)与(7.1.17)的解. ▮

由于约束问题(1.1.1)～(1.1.3)可写成

$$\min_{x\in\mathbf{R}^n} f(x), \tag{7.1.20}$$

$$\text{s.t.}\ \|\bar{c}(x)\|_2\leqslant 0. \tag{7.1.21}$$

所以, 当 $\delta=\|c(x(\sigma))\|_2$充分小时, 我们可把(7.1.16)与(7.1.17)看成是(7.1.20)与(7.1.21)的一个很好的近似, 从而可把 $x(\sigma)$当作(7.1.20) 与 (7.1.21) 的近似解. 罚函数法正是基于这一基本点, 每次迭代增加 σ_i 来逐步缩小$\|\bar{c}(x(\sigma))\|_2$. 下面给出的是利用 Courant 罚函数的罚函数法:

算法 7.1.3

步 1 给出 x_1, $\sigma_1>0$, $k:=1$; $\varepsilon\geqslant 0$.

步 2 利用初值 x_k 求解

$$\min f(x)+\sigma_k\|\bar{c}(x)\|_2^2, \tag{7.1.22}$$

得到解 $x(\sigma_k)$;

步 3 如果 $\|\bar{c}(x(\sigma_k))\|_2 \leqslant \varepsilon$, 则停;

$$x_{k+1}=x(\sigma_k),$$

$$\sigma_{k+1}=10\sigma_k,$$

$k:=k+1$; 转步 2.

我们有如下收敛性结果:

定理 7.1.4 设算法 7.1.3 中的误差界 ε 满足

$$\varepsilon > \min_{x\in \mathbf{R}^n}\|\bar{c}(x)\|_2, \tag{7.1.23}$$

则算法必有限终止.

证明 假定定理不真, 则必有 $\sigma_k\to+\infty$, 且对一切 k 均有

$$\|\bar{c}(x(\sigma_k))\|_2 \geqslant \varepsilon. \tag{7.1.24}$$

因为 ε 满足(7.1.23), 存在 $\hat{x}$, 使得

$$\|\bar{c}(\hat{x})\|_2 < \varepsilon. \tag{7.1.25}$$

由于 $x(\sigma_k)$的定义, 我们有

$$\begin{aligned} f(\hat{x})+\sigma_k\|\bar{c}(\hat{x})\|_2^2 &\geqslant f(x(\sigma_k))+\sigma_k\|\bar{c}(x(\sigma_k))\|_2^2 \\ &\geqslant f(x(\sigma_1))+\sigma_k\|\bar{c}(x(\sigma_k))\|_2^2. \end{aligned} \tag{7.1.26}$$

于是, 令 $\sigma_k\to+\infty$,

$$\begin{aligned} &\|\bar{c}(\hat{x})\|_2^2-\|\bar{c}(x(\sigma_k))\|_2^2 \\ &\geqslant \frac{1}{\sigma_k}[f(x(\sigma_1))-f(\hat{x})] \longrightarrow 0. \end{aligned} \tag{7.1.27}$$

这与(7.1.24)、(7.1.25)矛盾. 从而定理成立. ∎

从上面定理可知, 若原问题有可行点, 则对任何给定的误差允许 $\varepsilon>0$, 算法均有限终止于(7.1.16)与(7.1.17)的解, 且 $\delta\leqslant\varepsilon$.

定理 7.1.5 如果算法 7.1.3 不有限终止, 则必有

$$\varepsilon \leqslant \min_{x\in \mathbf{R}^n}\|\bar{c}(x)\|_2, \tag{7.1.28}$$

且

$$\lim_{k\to\infty}\|\bar{c}(x(\sigma_k))\|_2=\min_{x\in \mathbf{R}^n}\|\bar{c}(x)\|_2, \tag{7.1.29}$$

$\{x(\sigma_k)\}$的任何聚点 x^* 都是问题

$$\min_{x\in \mathbf{R}^n} f(x), \tag{7.1.30}$$

$$\text{s.t.} \|\bar{c}(x)\|_2 = \min_{x\in R^n} \|\bar{c}(x)\| \tag{7.1.31}$$

的解.

证明 由定理 7.1.4 知如果算法 7.1.3 不有限终止，则(7.1.28)成立.

由于 $\sigma_k \to +\infty$，与定理 7.1.4 的证明完全类似，我们可证(7.1.29)必成立.

设 x^* 是$\{x(\sigma_k)\}$的任一聚点,由(7.1.29)即知 x^* 必是(7.1.31)的可行点. 如果 x^* 不是(7.1.30)与(7.1.31)的解，则必存在 $\tilde{x}$，使得

$$f(\tilde{x}) < f(x^*), \tag{7.1.32}$$

且 $\tilde{x}$ 满足(7.1.31). 由于 $f(x(\sigma_k))$ 单调，且 x^* 是$\{x(\sigma_k)\}$的一个聚点，故必有

$$\lim_{k\to\infty} f(x(\sigma_k)) = f(x^*). \tag{7.1.33}$$

从(7.1.32)与(7.1.33)知:对一切充分大的 k 均有

$$f(\tilde{x}) < f(x(\sigma_k)). \tag{7.1.34}$$

由于 $\tilde{x}$ 是(7.1.31)的可行点,故

$$\|\bar{c}(\tilde{x})\|_2 \leqslant \|c(x(\sigma_k))\|_2 \tag{7.1.35}$$

对一切 k 都成立. 从(7.1.34)与(7.1.35)可推得

$$P(\tilde{x}, \sigma_k) < P(\tilde{x}(\sigma_k), \sigma_k). \tag{7.1.36}$$

这与 $x(\sigma_k)$的定义相矛盾. 所以定理为真. ▌

定理 7.1.5 的一个特殊情况是原问题有可行点时,如果 $\varepsilon = 0$，则算法 7.1.3 产生的点列 $\{x_k\}$ 的任何聚点都是原问题的解. 此时,我们有

$$\nabla f(x_{k+1}) + \sigma_k \sum_{i=1}^{m} \bar{c}_i(x_{k+1}) \nabla \bar{c}_i(x_{k+1}) = 0. \tag{7.1.37}$$

于是,定义

$$\lambda_i^{(k+1)} = -\sigma_k \bar{c}_i(x_{k+1}), \tag{7.1.38}$$

则 $\lambda^{(k)}$ 是 Lagrange 乘子 λ^* 的一个近似. 事实上,假定 $x_k \to x^*$，且 $\nabla c_i(x^*)$ $(i \in E \cup I(x^*))$线性无关,则 $\lambda^{(k)} \to \lambda^*$.

下面我们介绍两个内点罚函数. 一个是由 Carroll(1961) 提出的倒数罚函数:

$$P(x,\sigma)=f(x)+\sigma^{-1}\sum_{i=1}^{m}\frac{1}{c_i(x)};\qquad(7.1.39)$$

另一个是由 Frisch(1955)提出的对数罚函数:

$$P(x,\sigma)=f(x)-\sigma^{-1}\sum_{i=1}^{m}\log c_i(x).\qquad(7.1.40)$$

这两个罚函数均是用来处理不等式约束问题, 即 $m_e=0$. 罚函数(7.1.39)和(7.1.40)都是内点罚函数. 内点罚函数在可行域的边界无界, 故被称为*障碍函数*. 考虑极小化问题:

$$\min f(x)+\frac{1}{\sigma}\sum_{i=1}^{m}\frac{1}{c_i(x)},\qquad(7.1.41)$$

设给出的初始点 $\hat{x}$ 是严格内点, 即

$$c_i(x)>0,\ i=1,\ \cdots,\ m.\qquad(7.1.42)$$

则由于函数(7.1.39)在可行域

$$X=\{x\,|\,c_i(x)\geqslant 0,\ i=1,\ 2,\ \cdots,\ m\}\qquad(7.1.43)$$

的边界上无界, 故(7.1.41)的解 $x(\sigma)$ 必是可行点. 对于倒数罚函数, 我们有:

引理 7.1.6 记 $x(\sigma)$ 是 (7.1.41)在(7.1.43)的解, 对 $\sigma_2\geqslant\sigma_1>0$, 有:

$$f(x(\sigma_2))<f(x(\sigma_1)),\qquad(7.1.44)$$

$$\sum_{i=1}^{m}\frac{1}{c_i(x(\sigma_2))}\geqslant\sum_{i=1}^{m}\frac{1}{c_i(x(\sigma_1))}.\qquad(7.1.45)$$

证明 与引理 7.1.1 的证明完全类似. ∎

与算法 7.1.3 类似, 我们可构造利用倒数罚函数的内点罚函数法.

算法 7.1.7

步 1 给出内点 x_1, $\sigma_1>0$, $\varepsilon\geqslant 0$, $k:=1$.

步 2 利用初值 x_k 求解

$$\min f(x)+\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x)},\qquad(7.1.46)$$

得到解 $x(\sigma_k)$，令 $x_{k+1}=x(\sigma_k)$；

步3 如果

$$\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x_{k+1})}\leqslant\varepsilon, \tag{7.1.47}$$

则停；$\sigma_{k+1}=10\sigma_k$；$k:=k+1$；转步2.

与引理 7.1.2 类似，我们可证以下引理：

引理 7.1.8 设算法 7.1.8 有限终止于 x_{k+1}，则 x_{k+1} 是问题

$$\min_{x\in\mathbb{R}^n} f(x) \tag{7.1.48}$$

$$\text{s. t}\sum_{i=1}^{m}\frac{1}{c_i(x)}\leqslant\delta' \tag{7.1.49}$$

的解，其中 $\delta'=\sum_{i=1}^{m}1/c_i(x_{k+1})$.

证明 对任何满足(7.1.49)的 x，由 x_{k+1} 的定义可知

$$f(x)+\sigma_k^{-1}\sum_{i=1}^{m}\frac{1}{c_i(x)}\geqslant f(x_{k+1})+\sigma_k^{-1}\sum_{i=1}^{m}\frac{1}{c_i(x_{k+1})}. \tag{7.1.50}$$

利用(7.1.49)和(7.1.50)即知

$$f(x)\geqslant f(x_{k+1}). \tag{7.1.51}$$

所以定理成立. ▮

问题(7.1.48)与(7.1.49)可理解为问题

$$\min f(x), \tag{7.1.52}$$

$$\sum_{i=1}^{m}\frac{1}{c_i(x)}<+\infty \tag{7.1.53}$$

的一种近似，而(7.1.52)与(7.1.53)与不等式约束优化问题(1.1.1)～(1.1.3) $(m_e=0)$的差别在于可行域的边界是否为可行点. 当算法终止得到的点 x_{k+1} 十分靠近边界时，$\delta'=\sum_{i=1}^{m}1/c_i(x_{k+1})$ 将会非常大. 从而(7.1.49)的可行域将靠近原问题的可行域(7.1.43). 故我们可把 x_{k+1} 取为问题的近似解. 当算法终止点 x_{k+1} 不靠近边界时，这时，如果 ε 充分小，则知罚因子充分大. 从而在 x_{k+1} 点附近(7.1.46)的目标函数(即倒数罚函数)将与 $f(x)$

十分靠近．于是我们可把 x_{k+1} 近似地看成 $f(x)$ 的局部极小点．更为精确地，我们有以下定理：

定理 7.1.9 设算法 7.1.7 产生的点列非有限终止，则必有 $\varepsilon=0$，

$$\lim_{k\to\infty}\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x_{k+1})}=0, \tag{7.1.54}$$

且

$$\lim_{k\to\infty}f(x_k)=\inf_{x\in\text{int}(X)}f(x). \tag{7.1.55}$$

其中 $\text{int}(X)=\{x|c_i(x)>0,\ i=1,\cdots,m\}$．$\{x_k\}$ 的任何聚点 x^* 都是问题(1.1.1)～(1.1.3)($m_e=0$)的解．

证明 对任给的 $\eta>0$，存在向量 $x_\eta\in\text{int}(X)$，使得

$$f(x_\eta)\leqslant\inf_{x\in\text{int}(X)}f(x)+\eta/2. \tag{7.1.56}$$

由于 $\sigma_k\to+\infty$，故必存在 $\bar{k}$，当 $k\geqslant\bar{k}$ 时，

$$\sigma_k\geqslant\frac{2}{\eta}\sum_{i=1}^{m}\frac{1}{c_i(x_\eta)}. \tag{7.1.57}$$

于是，当 $k\geqslant\bar{k}$ 时，有

$$\begin{aligned}\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x_{k+1})}&\leqslant f(x_\eta)+\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x_\eta)}-f(x_{k+1})\\&\leqslant\inf_{x\in\text{int}(X)}f(x)+\frac{1}{2}\eta+\frac{1}{2}\eta-f(x_{k+1})\\&\leqslant\eta.\end{aligned} \tag{7.1.58}$$

由于 $\eta>0$ 的任意性，我们知(7.1.54)必成立．从终止条件(7.1.47)和极限(7.1.54)可知 $\varepsilon=0$．

由于 $\sigma_k\to0$，由 x_{k+1} 的定义和(7.1.56)可知

$$f(x_{k+1})\leqslant\inf_{x\in\text{int}(X)}f(x)+\frac{\eta}{2}+\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{c_i(x_\eta)}, \tag{7.1.59}$$

所以当 $k\geqslant\bar{k}_1$ 时，

$$f(x_{k+1})\leqslant\inf_{x\in\text{int}(X)}f(x)+\eta, \tag{7.1.60}$$

由 $\eta>0$ 的任意性即知(7.1.55)为真．从(7.1.55)可证 $\{x_k\}$ 的任何聚点都是问题(1.1.1)～(1.1.3)的解． ∎

对于倒数罚函数，我们有

$$\nabla f(x_{k+1})=\frac{1}{\sigma_k}\sum_{i=1}^{m}\frac{1}{[c_i(x_{k+1})]^2}\nabla c_i(x_{k+1}), \tag{7.1.61}$$

所以，近似的 Lagrange 乘子为

$$\lambda_i^{(k+1)}=\frac{1}{\sigma_k[c_i(x_{k+1})]^2}. \tag{7.1.62}$$

显然有

$$\lambda_i^{(k+1)}\geqslant 0, \tag{7.1.63}$$

且由(7.1.54)可知

$$\lambda_i^{(k+1)}c_i(x_{k+1})\to 0. \tag{7.1.64}$$

同样地，对于对数罚函数，我们也可构造类似的求解不等式约束优化的对数罚函数法，并可证明相似的收敛性结果。

罚函数法的收敛速度有如下估计：

$$\|x_k-x^*\|_2=O(1/\sigma_k), \tag{7.1.65}$$

即在假定收敛的前提下，误差$\|x_k-x^*\|_2$和$1/\sigma_k$是一个量级的（见 Fletcher，1987）。而且(7.1.65)是不能改进的。现在考虑下面例子：

例 7.1.10

$$\min_{(t,u)\in\mathbb{R}^2} t+u, \tag{7.1.66}$$

$$\text{s. t.}\ u-t^2=0. \tag{7.1.67}$$

如果我们用极小化简单罚函数

$$\min t+u+\sigma[u-t^2]^2 \tag{7.1.68}$$

求得$x(\sigma)$，则知

$$x(\sigma)=\begin{bmatrix}-\frac{1}{2}\\ \frac{1}{4}-\frac{1}{2\sigma}\end{bmatrix}=x^*-\begin{bmatrix}0\\ \frac{1}{2}\end{bmatrix}\frac{1}{\sigma}, \tag{7.1.69}$$

其中$x^*=\left(-\frac{1}{2},\ \frac{1}{4}\right)^T$是问题(7.1.66)与(7.1.67)的唯一解。

由(7.1.69)即知(7.1.65)不可能改进。于是，我们需要解x_k达到精度ε，则要求

$$\sigma_k = O(1/\varepsilon). \tag{7.1.70}$$

所以，由罚函数法求解约束优化问题往往需要很大的罚因子．从而使求解罚函数的极小时，在数值上有困难.

§7.2 乘子罚函数

早期罚函数的缺点是需要罚因子趋于无穷大，才可使求罚函数极小和求解原问题等价．乘子罚函数利用近似 Lagrange 乘子，从而不需要无穷大的罚因子．

为了叙述简单，我们考虑等式约束问题

$$\min_{x\in R^n} f(x) \tag{7.2.1}$$

$$\text{s. t.} c_i(x)=0,\ i=1,\ \cdots,\ m. \tag{7.2.2}$$

我们记 $c(x)=(c_1(x),\ \cdots,\ c_m(x))^T$; $A(x)=[\nabla c(x)]^T$, 设 x^* 是问题(7.2.1)与(7.2.2)的解，且 λ^* 是相应的 Lagrange 乘子．由 Kuhn-Tucker 定理知 x^* 必是 Lagrange 函数

$$L(x,\ \lambda^*)=f(x)-(\lambda^*)^Tc(x) \tag{7.2.3}$$

的稳定点．但一般说来，x^* 并不是 Lagrange 函数(7.2.3)的极小点，故我们考虑

$$P(x,\ \lambda^*,\ \sigma)=L(x,\ \lambda^*)+\frac{1}{2}\sigma\|c(x)\|_2^2, \tag{7.2.4}$$

由于 $c(x^*)=0$, 不难发现

$$\nabla_x P(x^*,\ \lambda^*,\ \sigma)=0, \tag{7.2.5}$$

$$\nabla_{xx}^2 P(x^*,\ \lambda^*,\ \sigma)=\nabla_{xx}^2 L(x^*,\ \lambda^*)+\sigma A(x^*)A(x^*)^T. \tag{7.2.6}$$

设在 x^* 处二阶充分条件满足，即对一切满足 $A(x^*)^Td=0$ 的非零向量，均有

$$d^T\nabla_{xx}^2 L(x^*,\ \lambda^*)d>0\ . \tag{7.2.7}$$

在二阶充分条件的假定下，我们可证必存在 σ_1, 使得当 $\sigma\geqslant\sigma_1$ 时，

$$\nabla_{xx}^2 L(x^*,\ \lambda^*)+\sigma A(x^*)A(x^*)^T \tag{7.2.8}$$

是正定阵．于是对于充分大的 σ（只要 $\sigma>\sigma_1$），x^* 也是函数(7.2.4)的极小点．由于我们事先并不知道 λ^*，故用乘子 λ 代替，

从而得到增广(augmented)Lagrange 罚函数:

$$P(x, \lambda, \sigma)=f(x)-\lambda^T c(x)+\frac{1}{2}\sigma\|c(x)\|_2^2. \qquad (7.2.9)$$

这个罚函数最早是由 Henstenes(1969)导出的.

增广 Lagrange 函数 (7.2.9) 的一种等价形式是由 Powell (1969) 独立提出的. Powell(1969) 的思想非常简单, 考虑简单罚函数

$$P(x, \sigma)=f(x)+\frac{1}{2}\sigma\|c(x)\|_2^2, \qquad (7.2.10)$$

要求 $\nabla_x P(x, \sigma)=0$, 则得到

$$\nabla f(x)+\sigma\sum_{i=1}^{m} c_i(x)\nabla c_i(x)=0. \qquad (7.2.11)$$

由于我们要求 $c_i(x)\to 0$, $\sigma c_i(x)\to\lambda_i^*$, 故必然导致 $\sigma\to+\infty$. 于是 Powell(1969)提出对 $c_i(x)$进行平移, 即用 $c_i(x)-\theta_i$ 代替 $c_i(x)$, θ_i 是参数. 这种平移的好处是不破坏 $\nabla c_i(x)$ 的方向. 由此, Powell(1969)得到罚函数

$$P(x, \theta, \sigma)=f(x)+\frac{1}{2}\sigma\sum_{i=1}^{m}(c_i(x)-\theta_i)^2. \qquad (7.2.12)$$

如果我们定义 $\lambda_i=\sigma\theta_i$, 则知 (7.2.12)和(7.2.9) 只相差与 x 无关的项 $\frac{1}{2}\sigma\sum_{i=1}^{m}\theta_i^2$. 正由于这种等价性, 罚函数 (7.2.9) 也称为 Henstenes-Powell罚函数.

对于一般约束规划, Rockafellar(1973)将(7.2.9) 推广到如下罚函数:

$$P(x, \lambda, \sigma)=f(x)-\sum_{i=1}^{m_e}\left[\lambda_i c_i(x)-\frac{1}{2}\sigma_i(c_i(x))^2\right]$$
$$-\sum_{i=m_e+1}^{m}\begin{cases}\lambda_i c_i(x)-\frac{1}{2}\sigma_i(c_i(x))^2, & \text{如果 } c_i(x)<\lambda_i/\sigma_i;\\ \frac{1}{2}\lambda_i^2/\sigma_i, & \text{否则.}\end{cases} \qquad (7.2.13)$$

事实上, Rockafellar 考虑的是特殊情形 $m_e=0$ 以及 $\sigma_i\equiv\sigma$.

(7.2.13)也可以写成下列形式(相差一个与 x 无关的项):

$$P(x,\ \theta,\ \sigma)=f(x)+\frac{1}{2}\sum_{i=1}^{m_e}\sigma_i[c_i(x)-\theta_i]^2$$

$$+\frac{1}{2}\sum_{i=m_e+1}^{m}\sigma_i(c_i(x)-\theta_i)_-^2. \qquad (7.2.14)$$

上式是由 Fletcher(1975)导出的，它显然是罚函数(7.2.12)的直接推广.

假定我们给定 $\lambda_i^{(k)}$、$\sigma_i^{(k)}$，求得

$$\min P(x,\ \lambda^{(k)},\ \sigma^{(k)}) \qquad (7.2.15)$$

的解为 $\bar{x}_k$. 其中(7.2.15)中的目标函数是罚函数(7.2.13). 我们必有

$$\nabla f(\bar{x}_k)=\sum_{i=1}^{m_e}[\lambda_i^{(k)}-\sigma_i^{(k)}c_i(\bar{x}_k)]\ \nabla c_i(\bar{x}_k)$$

$$+\sum_{i=m_e+1}^{m}\max\{\lambda_i^{(k)}-\sigma_i^{(k)}c_i(\bar{x}_k),\ 0\}\nabla c_i(\bar{x}_k). \qquad (7.2.16)$$

于是，我们可取

$$\lambda_i^{(k+1)}=\lambda_i^{(k)}-\sigma_i^{(k)}c_i(\bar{x}_k),\ i=1,\ \cdots,\ m_e, \qquad (7.2.17)$$

$$\lambda_i^{(k+1)}=\max\{\lambda_i^{(k)}-\sigma_i^{(k)}c_i(\bar{x}_k),\ 0\},\ i=m_e+1,\ \cdots,\ m. \qquad (7.2.18)$$

为下次迭代的 Lagrange 乘子. 基于增广 Lagrange 函数的罚函数方法可写成下面形式:

算法 7.2.1

步 1 给出初始 x_1; 并给出 $\lambda^{(1)}$、$\sigma^{(1)}>0$, 且 $\lambda_i^{(1)}\geqslant 0(i\in I)$; $\varepsilon\geqslant 0$, $k:=1$.

步 2 求解(7.2.15), 给出 $\bar{x}_k$;

$$x_{k+1}:=\bar{x}_k;$$

如果 $\|\bar{c}(x_{k+1})\|\leqslant\varepsilon$, 则停;

步 3 如果 $\|\bar{c}(x_{k+1})\|_2<\frac{1}{4}\|\bar{c}(x_k)\|_2$, 则转步 4;

$\sigma^{(k)}:=10\sigma^{(k)}$; 转步 2;

步 4 由(7.2.17)~(7.2.18), 计算 $\lambda^{(k+1)}$;

$\sigma^{(k+1)}:=\sigma^{(k)}$; $k:=k+1$; 转步 2.

在上面算法中, $\bar{c}(x)$是由(7.1.4)定义的.

由算法 7.2.1 的收敛性, 不难建立以下定理:

定理 7.2.2 如果问题(1.1.1)～(1.1.3)有可行点, 则对任何 $\varepsilon>0$, 算法 7.2.1 必有限终止.

证明 假定算法 7.2.1 不有限终止, 则存在 $\bar{k}$, 使得

$$\|\bar{c}(x(\hat{\sigma}_j))\|_2>\frac{1}{4}\|\bar{c}(x_{\bar{k}})\|_2. \tag{7.2.19}$$

其中 $x(\hat{\sigma}_j)$是

$$\min P(x,\ \lambda^{(\bar{k})},\ \hat{\sigma}_j) \tag{7.2.20}$$

的极小点, $\hat{\sigma}_j(j=1,\ 2,\ \cdots)$是一趋于正无穷的数列. 由于 $x(\hat{\sigma}_j)$是(7.2.20)的解, 故有

$$\begin{aligned}&f(x(\hat{\sigma}_j))-\sum_{i=1}^{m_e}\left[\lambda_i^{(\bar{k})}c_i(x(\hat{\sigma}_i))-\frac{1}{2}(\hat{\sigma}_j)_i^2[c_i(x(\hat{\sigma}_j))]^2\right]\\&\quad+\sum_{i=m_e+1}^{m}\frac{1}{2}(\hat{\sigma}_j)_i\left\{\left[c_i(x(\hat{\sigma}_j))-\frac{\lambda_i^{(\bar{k})}}{(\hat{\sigma}_j)_i}\right]^2-\left[\frac{\lambda_i^{(\bar{k})}}{(\hat{\sigma}_j)_i}\right]^2\right\}\\&\leqslant f(\bar{x}).\end{aligned} \tag{7.2.21}$$

其中 $\bar{x}$ 是(1.1.1)～(1.1.3)的任一可行点. 与引理 7.1.1 类似, 我们可证对充分大的 j, 有

$$f(x(\hat{\sigma}_{j+1}))\geqslant f(x(\hat{\sigma}_j)). \tag{7.2.22}$$

于是在(7.2.21)中令 $\hat{\sigma}_j\to+\infty$, 即得到

$$\|\bar{c}(x(\hat{\sigma}_j))\|_2\to 0. \tag{7.2.23}$$

这显然与算法不有限终止相矛盾. 所以定理成立. ▌

定理 7.2.3 设问题 (1.1.1)～(1.1.3) 有可行点, 则算法 7.2.1 产生的点列$\{x_k\}$的任何聚点 x^* 都是可行点, 如果 $\lambda^{(k)}$ 有界, 则 x^* 必是原问题(1.1.1)～(1.1.3)的解.

证明 从定理 7.2.2 的证明可知 x^* 必是可行点. 假定 $\lambda^{(k)}$ 对一切 k 一致有界, 由(7.2.21)式即知

$$f(x^*)\leqslant f(\bar{x}). \tag{7.2.24}$$

由于 $\bar{x}$ 的任意性, 即知 x^* 是(1.1.1)～(1.1.3)的解. ▌

下面考虑算法 7.2.1 的收敛速度. 为了叙述简单起见, 不妨设

仅有等式约束. 我们假定 $x_k \to x^*$, 且 $A(x^*)=\nabla c(x^*)^T$ 是列满秩. 不失一般性, 可假定所有的 $A(x_k)=\nabla c(x_k)^T$ 都是列满秩的. 于是算法产生的 $\lambda^{(k+1)}$ 和下式

$$A(x_{k+1})\lambda(x_{k+1})=g(x_{k+1}) \tag{7.2.25}$$

定义的 $\lambda(x_{k+1})$ 是完全等价的, 其中 $g(x)=\nabla f(x)$. 考虑函数

$$\lambda(x)=[A(x)]^+g(x), \tag{7.2.26}$$

不难求得

$$[\nabla\lambda(x)^T]^T=[A(x)]^+w(x), \tag{7.2.27}$$

其中

$$w(x)=\nabla^2 f(x)-\sum_{i=1}^m[\lambda(x)]_i\nabla^2 c_i(x). \tag{7.2.28}$$

由算法知

$$\lambda(x_{k+1})+D_k c(x_{k+1})=\lambda(x_k). \tag{7.2.29}$$

其中

$$D_k=\begin{bmatrix}\sigma_1^{(k)} & & 0\\ & \ddots & \\ 0 & & \sigma_m^{(k)}\end{bmatrix}. \tag{7.2.30}$$

于是, 我们有

$$\begin{aligned}&[D_k A(x^*)^T+A(x^*)^+w(x^*)](x_{k+1}-x^*)\\ &\approx A(x^*)^+w(x^*)(x_k-x^*).\end{aligned} \tag{7.2.31}$$

从而可知, 除非 $\sigma^{(k)}\to+\infty$, 算法 7.2.1 产生的点列一般是线性收敛的.

增广 Lagrange 函数 (7.2.13) 的一个缺点是: 它仅是一次连续可微函数, 所以在求解(7.2.15)时, 有可能在数值上有困难.

Fletcher(1973)提出对于等式约束问题的第一个光滑精确罚函数(smooth exact penalty function)为

$$P(x,\sigma)=f(x)-\lambda(x)^T c(x)+\frac{1}{2}c(x)^T D c(x). \tag{7.2.32}$$

其中

$$\lambda(x)=(A(x))^+g(x) \tag{7.2.33}$$

$$D=\begin{bmatrix}\sigma_1 & & 0\\ & \ddots & \\ 0 & & \sigma_m\end{bmatrix}. \qquad (7.2.34)$$

显然，设 x^* 是等式约束问题(7.2.1)与(7.2.2)的解，且 $A(x^*)$ 列满秩，我们有

$$\nabla_x P(x^*,\ \sigma)=g(x^*)-A(x^*)\lambda(x^*)=0, \qquad (7.2.35)$$

$$\nabla^2_{xx} P(x^*,\ \sigma)=w(x^*)+A(x^*)D(A(x^*))^T, \qquad (7.2.36)$$

其中

$$w(x^*)=\nabla^2_{xx}L(x^*,\ \lambda(x^*)). \qquad (7.2.37)$$

与 (7.2.8) 类似，在二阶充分条件的假定下，我们可证必存在 $\bar{\sigma}>0$，使得当 $\sigma_i\geqslant\bar{\sigma}$ 时，矩阵(7.2.36)为正定阵．于是我们可知当 $\sigma_i\geqslant\bar{\sigma}$ 时，x^* 也是(7.2.32)的严格极小点．在(7.2.32)中令所有的 σ_i 相等，则得到简单形式的 Fletcher 光滑精确罚函数：

$$P(x,\ \sigma)=f(x)-\lambda(x)^Tc(x)+\frac{1}{2}\sigma\|c(x)\|_2^2, \qquad (7.2.38)$$

由于(7.2.38)是精确罚函数，故只要 σ_i 充分大时，求解(7.2.32)的极小 $\bar{x}$ 且有 $c(\bar{x})=0$ 即知 $\bar{x}$ 也是等式约束问题(7.2.1)与(7.2.2)的解．

精确罚函数(7．2.32)的优点是它光滑，即多次连续可微．这样无约束优化问题

$$\min_{x\in\mathbf{R}^n} P(x,\ \sigma) \qquad (7.2.39)$$

的求解方法的收敛速度可以很快(如拟牛顿法)．而我们知道早期罚函数方法要求乘子无限增大使求罚函数极小很困难；增广 Lagrange 函数的罚函数方法也可能使求解(7.2.15)有困难，这是由于(7.2.13)的二阶导数有间断点．罚函数(7.2.32)的一个美中不足的地方是，计算 $\nabla_x P(x,\ \sigma)$时需要计算 $\nabla\lambda(x)^T$．由(7.2.27)与(7.2.28)可知，我们需要计算 $\nabla^2 f(x)$和$\nabla^2 c_i(x)$ $(i=1,\ \cdots,\ m)$．这样，我们在利用导数的方法求解(7.2.39)时，必须计算 $\nabla^2 f(x)$、$\nabla^2 c_i(x)$ $(i=1,\ \cdots,\ m)$，从而不仅要求 $\nabla^2 f(x)$、$\nabla^2 c_i(x)$ $(i=1,\ \cdots,\ m)$可计算，而且使算法的计算量非常大．

光滑精确罚函数不仅提供了光滑函数等价地描述约束优化问题，而且在不少约束优化的计算方法中，有直接应用。关于光滑精确罚函数的其他讨论可见 Bertsekas(1980, 1982a, b)、Boggs 和 Tolle(1981)、Di Pillo 和 Grippo(1979)以及 Han 和 Mangasarian (1983)等。

§7.3 非光滑精确罚函数

最常用的精确罚函数是 L_1 精确罚函数：

$$P(x, \sigma)=f(x)+\sigma\left[\sum_{i=1}^{m_e}|c_i(x)|+\sum_{i=m_e+1}^{m}|c_i(x)_-|\right]. \qquad (7.3.1)$$

(7.3.1) 最初是由 Zangwill(1967b) 提出的，它和简单罚函数(7.1.2)的不同之处在于：将约束违反度的平方换成绝对值。利用记号(7.1.4)，可将(7.3.1)写成

$$P(x, \sigma)=f(x)+\sigma\|\bar{c}(x)\|_1. \qquad (7.3.2)$$

设 x^* 是(1.1.1)～(1.1.3)的 Kuhn-Tucker 点，如果 $\sigma>\|\lambda^*\|_\infty$，则可证

$$f(x^*)+d^T\nabla f(x^*)+\sigma\left[\sum_{i=1}^{m_e}|c_i(x^*)+d^T\nabla c_i(x^*)|\right.$$
$$\left.+\sum_{i=m_e+1}^{m}|(c_i(x^*)+d^T\nabla c_i(x^*))_-|\right]\geqslant P(x^*, \sigma) \qquad (7.3.3)$$

对一切充分小的 d 都成立，所以 x^* 也是(7.3.2)的一阶稳定点。如果二阶充分条件满足，则有以下定理：

定理 7.3.1 设 x^* 是(1.1.1)～(1.1.3)的局部极小点，λ^* 是相应的 Lagrange 乘子。如果二阶充分条件满足，且 $\sigma>\|\lambda^*\|_\infty$，则 x^* 是(7.3.2)的局部严格极小点。

证明 由于二阶充分条件满足，类似(7.2.8)我们可证存在 σ_1，使得 x^* 是函数

$$L(x, \lambda^*)+\frac{1}{2}\sigma_1\|\bar{c}(x)\|_2^2 \qquad (7.3.4)$$

的局部严格极小点。由于 $\sigma>\|\lambda^*\|_\infty$ 和 $\|\bar{c}(x^*)\|=0$，存在 $\delta>0$，使

得 $\|x-x^*\|\leqslant\delta$ 时，有

$$\sigma_1\|\bar{c}(x)\|_2<\sigma-\|\lambda^*\|_\infty. \tag{7.3.5}$$

从而，当 $\|x-x^*\|\leqslant\delta$ 时，有

$$\begin{aligned} P(x,\ \sigma) &= f(x)+\|\lambda^*\|_\infty\|\bar{c}(x)\|_1+(\sigma-\|\lambda^*\|_\infty)\|\bar{c}(x)\|_1 \\ &\geqslant L(x,\ \lambda^*)+\sigma_1\|\bar{c}(x)\|_1\|\bar{c}(x)\|_2 \\ &\geqslant L(x,\ \lambda^*)+\frac{1}{2}\sigma_1\|\bar{c}(x)\|_2^2 \\ &> L(x^*,\ \lambda^*)=P(x^*,\ \sigma). \end{aligned} \tag{7.3.6}$$

所以定理为真. ▮

所以只要有 $\sigma>\|\lambda^*\|_\infty$，则只需求解一个无约束优化问题.

$$\min_{x\in\mathbb{R}^n} P(x,\ \sigma), \tag{7.3.7}$$

即解决了(1.1.1)～(1.1.3). 但不幸的是，罚函数(7.3.1)是一个不可微函数，而且在一般情况下，解 x^* 是不可微的. 从而一些收敛较快的利用导数的方法不能直接用来求解(7.3.7).

第 8 章

线性约束规划

比二次规划稍复杂的一类约束优化问题是线性约束优化问题，即在线性约束下求解一般的非线性函数的极小值：

$$\min_{x\in \mathbf{R}^n} f(x), \tag{8.0.1}$$

$$\text{s.t.}\, a_i^T x = b_i,\ i=1,\ \cdots,\ m_e; \tag{8.0.2}$$

$$a_i^T x \geqslant b_i,\ i=m_e+1,\ \cdots,\ m, \tag{8.0.3}$$

其中 $f(x)$是 $\mathbf{R}^n$ 上的非线性函数．和二次规划一样，在线性约束优化问题中，有一些约束可写成(6.0.4)或(6.0.5)的形式．由于线性约束规划问题的约束条件和二次规划的约束条件具有同样的形式，所以处理线性约束规划的约束的方法和处理二次规划的约束时是一样的．因为 $f(x)$一般不是二次函数，所以求解(8.0.1)～(8.0.3)比解二次规划要困难，而且算法一般不会像二次规划的方法那样具有二次终止性．

§8.1 等式约束

等式线性约束规划问题是指 $m_e=m$ 这一特殊情形下的线性约束规划问题(8.0.1)～(8.0.3)．它可写成

$$\min_{x\in \mathbf{R}^n} f(x), \tag{8.1.1}$$

$$\text{s.t.}\, A^T x = b, \tag{8.1.2}$$

其中 $A\in \mathbf{R}^{n\times m}$，$b\in \mathbf{R}^m$．与 6.2 节类似，我们假定秩$(A)=m$．

下面将 6.2 节讨论的广义消去法推广到问题(8.1.1)与(8.1.2)．设矩阵 $Y\in \mathbf{R}^{n\times m}$，$Z\in \mathbf{R}^{n\times (n-m)}$ 满足秩$(Z)=n-m$ 且

$$A^TY=I,\quad A^TZ=0. \tag{8.1.3}$$

则知(8.1.2)的所有可行点可表示为

$$x=Yb+Z\hat{x},\quad \hat{x}\in\mathbf{R}^{n-m}. \tag{8.1.4}$$

于是(8.1.1)与(8.1.2)等价于无约束优化问题:

$$\min_{\hat{x}\in\mathbf{R}^{n-m}} f(Yb+Z\hat{x}). \tag{8.1.5}$$

定义 $\hat{f}(\hat{x})=f(Yb+Z\hat{x})$，我们先考虑(8.1.5)的最速下降法：设有当前迭代点 $\hat{x}_k$，则线搜索方向为

$$\hat{d}_k=-\nabla\hat{f}(\hat{x}_k)=-Z^T\nabla f(x_k), \tag{8.1.6}$$

其中 $x_k=Yb+Z\hat{x}_k$. 计算步长 $\hat{\alpha}_k$:

$$\hat{f}(\hat{x}_k+\hat{\alpha}_k\hat{d}_k)=\min_{\alpha>0}\hat{f}(\hat{x}_k+\alpha\hat{d}_k), \tag{8.1.7}$$

然后令 $\hat{x}_{k+1}=\hat{x}_k+\hat{\alpha}_k\hat{d}_k$. 定义

$$d_k=Z\hat{d}_k=-ZZ^T\nabla f(x_k), \tag{8.1.8}$$

则知 $\hat{\alpha}_k$ 也是问题

$$\min_{\alpha>0} f(x_k+\alpha d_k) \tag{8.1.9}$$

的解，所以 $x_{k+1}=Yb+Z\hat{x}_{k+1}$ 满足

$$x_{k+1}=x_k+\alpha_k d_k, \tag{8.1.10}$$

其中 $\alpha_k=\hat{\alpha}_k$ 是(8.1.9)的解. 从而我们得到了一个基于(8.1.5)的最速下降法如下:

算法 8.1.1

步 1　给出 x_1 满足(8.1.2)；给出 Z，满足 $A^TZ=0$ 且秩 $(Z)=n-m$; $k=1$, $\varepsilon\geqslant 0$;

步 2　计算(8.1.8)；如果 $\|d_k\|\leqslant\varepsilon$，则停;

步 3　求解(8.1.9)，得到 α_k;

$$x_{k+1}=x_k+\alpha_k d_k;$$

$k:=k+1$; 转步 2.

由于算法 8.1.1 实质上就是 $\min\hat{f}(\hat{x})$ 的最速下降法，从定理 3.1.2 可知：如果 $f(x_k)$ 不趋于 $-\infty$，则必有

$$\lim_{k\to\infty}\|Z^T\nabla f(x_k)\|=0. \tag{8.1.11}$$

于是对$\{x_k\}$的任何聚点x^*，均有

$$Z^T\nabla f(x^*)=0. \tag{8.1.12}$$

利用(8.1.12)、(8.1.3)以及秩$(A)=m$可证存在$\lambda^*\in\mathbf{R}^m$，使得

$$\nabla f(x^*)=A\lambda^*. \tag{8.1.13}$$

所以算法 8.1.1 产生点列的任何聚点都是问题(8.1.1)与(8.1.2)的 Kuhn-Tucker 点.

由 Z 的定义可知，Z^T 是从 $\mathbf{R}^n$ 到 A^T 的零空间 Null(A^T) 的一个映照，所以我们称 $Z^T\nabla f(x)$ 是 $f(x)$ 的既约梯度 (reduced gradient). 从而可称算法 8.1.1 为既约最速下降法.

从第 6.2 节的讨论可知，Z 的选取有很多种不同的形式. 如果 Z 由(6.2.29)给出，则有

$$ZZ^T=I-Q_1Q_1^T=I-AA^+, \tag{8.1.14}$$

其中 $A^+=(A^TA)^{-1}A^T$ 是 A 的广义逆矩阵. 在这种特殊情况下 $(ZZ^T)^2=ZZ^T$，故 ZZ^T 为一投影算子 (projection operator). 这时

$$d_k=-(I-AA^+)\nabla f(x_k). \tag{8.1.15}$$

算法 8.1.1 就是 Rosen(1960) 提出的梯度投影法 (gradient projection method). $(I-AA^+)\nabla f(x)$被称为投影梯度 (projected gradient).

如果函数 $f(x)$的二阶导数可计算，则利用公式

$$\nabla^2\hat{f}(\hat{x})=Z^T\nabla^2 f(x)Z \tag{8.1.16}$$

求出 $\hat{f}(\hat{x})$的二阶导数. 从而可利用牛顿法求解无约束优化问题(8.1.5). 由于矩阵(8.1.6)不一定是正定的，故我们求解带信赖域约束的牛顿步：

$$\min_{\hat{d}\in\mathbf{R}^{n-m}} \hat{g}_k^T\hat{d}+\frac{1}{2}\hat{d}^TZ^T\nabla^2 f(x_k)Z\hat{d}, \tag{8.1.17}$$

$$\text{s.t.}\ \|\hat{d}\|\leqslant\Delta_k, \tag{8.1.18}$$

其中 $\Delta_k>0$ 是第 k 次迭代的信赖域半径. $\hat{g}_k$ 是在 x_k 处的既约梯度：

$$\hat{g}_k=\hat{g}(x_k)=Z^Tg(x_k)=Z^T\nabla f(x_k). \tag{8.1.19}$$

于是我们可给出一个带信赖域技巧的牛顿法如下：

算法 8.1.2

步 1　给出 x_1 满足(8.1.2)，且给出 Z 满足 $A^TZ=0$，且秩$(Z)=n-m$。给出 $\Delta_1>0$，$\varepsilon\geqslant 0$，$k:=1$。

步 2　计算 $\hat{g}_k$；如果 $\|\hat{g}_k\|<\varepsilon$，则停；求解(8.1.17)与(8.1.18)得到 $\hat{d}_k$，$d_k=Z\hat{d}_k$；

步 3　计算

$$r_k=\frac{f(x_k)-f(x_k+d_k)}{-\hat{g}_k^T\hat{d}_k-\frac{1}{2}\hat{d}_k^TZ^T\nabla^2 f(x_k)Z\hat{d}_k}; \tag{8.1.20}$$

步 4　令

$$x_{k+1}=\begin{cases} x_k+d_k, & 如果\ r_k>0, \\ x_k, & 否则. \end{cases} \tag{8.1.21}$$

$$\Delta_{k+1}=\begin{cases} \frac{1}{2}\|\hat{d}_k\|, & 如果\ r_k<0.25; \\ 2\Delta_k, & 如果\ r_k>0.75\ 且\ \Delta_k=\|\hat{d}_k\|; \\ \Delta_k, & 其他情况. \end{cases} \tag{8.1.22}$$

步 5　$k:=k+1$；转步 2.

算法 8.1.2 是我们要在第10章第 10.3 节中讨论的算法 10.3.1 的特殊情形．于是我们知道只要$\{x_k\}$有界，则$\{x_k\}$必存在一个聚点 x^* 为问题(8.1.1)与(8.1.2)的 Kuhn-Tucker 点．假定 x_k 收敛于 x^*，且 $Z^T\nabla^2 f(x^*)Z$ 正定，由(8.1.20)知

$$r_k\to 1. \tag{8.1.23}$$

故知存在 $\bar{k}$，使得 $\Delta_k=\Delta_{\bar{k}}$ 对一切 $k\geqslant\bar{k}$ 都成立．于是对所有充分大的 k，有

$$\hat{x}_{k+1}=\hat{x}_k-(Z^T\nabla^2 f(x_k)Z)^{-1}Z^T\nabla f(x_k). \tag{8.1.24}$$

于是

$$\|\hat{x}_{k+1}-\hat{x}^*\|=O(\|\hat{x}_k-\hat{x}^*\|^2). \tag{8.1.25}$$

由于 $\hat{x}_k-\hat{x}^*=Z(x_k-x^*)$，$Z$ 为列满秩，我们从(8.1.25)可证

$$\|x_{k+1}-x^*\|=O(\|x_k-x^*\|^2). \tag{8.1.26}$$

所以算法 8.1.2 是平方收敛的．

如果不计算或者是不能计算二阶导数 $\nabla^2 f(x)$ 的话，一种办法是用有限差商

$$Z^T[g(x_k+z_i h)-g(x_k)]/h \tag{8.1.27}$$

来代替 $Z^T\nabla^2 f(x_k)Z$ 的第 i 列，其中 z_i 是 Z 的第 i 列，$h>0$ 是差商步长．另一种办法是用拟牛顿公式修正矩阵 B_k，而用 B_k 代替 $\nabla^2 f(x_k)$．考虑 BFGS 修正公式：

$$B_{k+1}=B_k-\frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}+\frac{y_k y_k^T}{s_k^T y_k}. \tag{8.1.28}$$

其中

$$s_k=x_{k+1}-x_k, \tag{8.1.29}$$

$$y_k=\nabla f(x_{k+1})-\nabla f(x_k). \tag{8.1.30}$$

考虑变换(8.1.4)，令 $\hat{B}_k=Z^T B_k Z$，则不难证明由(8.1.28)可推出

$$\hat{B}_{k+1}=\hat{B}_k-\frac{\hat{B}_k\hat{s}_k\hat{s}_k^T\hat{B}_k}{\hat{s}_k^T\hat{B}_k\hat{s}_k}+\frac{\hat{y}_k\hat{y}_k^T}{\hat{s}_k^T\hat{y}_k}. \tag{8.1.31}$$

其中

$$\hat{s}_k=\hat{x}_{k+1}-\hat{x}_k=(Z^T Z)^{-1}Z^T s_k, \tag{8.1.32}$$

$$\hat{y}_k=\nabla\hat{f}(\hat{x}_{k+1})-\nabla\hat{f}(\hat{x}_k)=Z^T y_k. \tag{8.1.33}$$

于是我们发现用拟牛顿法修正 $\nabla^2 f(x)$ 和用拟牛顿法修正 $\nabla^2\hat{f}(\hat{x})$ 是等价的．修正 $\hat{B}_k$ 的好处是 $\hat{B}_k\in\mathbf{R}^{(n-m)\times(n-m)}$ 当 $n-m\ll n$ 时，修正 $\hat{B}_k$ 比修正 B_k 计算量要小得多．下面给出的是一个求解(8.1.1)与(8.1.2)的拟牛顿法：

算法 8.1.3

步 1　给出 x_1 满足(8.1.2)，且给出 $\hat{B}_1$ 正定；给出 Z 满足 $A^T Z=0$ 且秩$(Z)=n-m$；$\varepsilon\geq 0$，$k:=1$；

步 2　计算 $\hat{g}_k=Z^T\nabla f(x_k)$；

如果 $\|\hat{g}_k\|\leq\varepsilon$，则停；

令：$\hat{d}_k=-\hat{B}_k^{-1}\hat{g}_k$

$d_k:=Z\hat{d}_k$.

步 3　利用线搜索方法求出 $\alpha_k>0$；

$x_{k+1}=x_k+\alpha_k d_k$，

$\hat{s}_k=\alpha_k\hat{d}_k$,

$\hat{y}_k=Z^T[\nabla f(x_{k+1})-\nabla f(x_k)]$,

由(8.1.31)计算 $\hat{B}_{k+1}$;

步4 $k:=k+1$; 转步2.

算法 8.1.3 实质上就是对于 $\hat{f}(\hat{x})$ 求极小的 BFGS 方法. 所以由第四章的结果可知如果 $f(x)$ 在子空间 $\{x \mid A^Tx=0\}$ 上是凸的, 则有

$$\liminf_{k\to\infty}\|Z^T\nabla f(x_k)\|=0, \tag{8.1.34}$$

从而 $\{x_k\}$ 存在聚点 x^* 使得 $Z^T\nabla f(x^*)=0$, 即 x^* 为问题(8.1.1)与(8.1.2)的 Kuhn-Tucker 点.

由于 Y 满足(8.1.3), 故由

$$g^*=A\lambda^* \tag{8.1.35}$$

可知

$$\lambda^*=Y^Tg^*, \tag{8.1.36}$$

于是

$$\lambda_k=Y^Tg(x_k) \tag{8.1.37}$$

是一个 Lagrange 乘子的估计值. 如果 Y 由(6.2.28)定义, 则有

$$\lambda_k=A^+g(x_k). \tag{8.1.38}$$

它是最小二乘 $\min\|A\lambda-g(x_k)\|_2^2$ 的最小范数解.

§8.2 积极集法

与 6.3 节类似, 我们可通过解一串等式约束问题来达到求解不等式约束问题(8.0.1)～(8.0.3)的目的.

积极集法是一个可行点方法. 其基本思想是: 如果 x_k 是等式约束问题

$$\min f(x) \tag{8.2.1}$$

$$a_i^Tx=b_i, \quad i\in S_k \tag{8.2.2}$$

的解, 其中

$$S_k=E\cup I_k \tag{8.2.3}$$

$E=\{1,\cdots,m_e\}$, $I_k\subseteq I=\{m_e+1,\cdots,m\}$, 则只要相应的 Lagrange 乘子 λ_k 满足

$$(\lambda_k)_i\geqslant 0,\quad i\in I_k. \tag{8.2.4}$$

x_k 就也是不等式约束问题 (8.0.1)~(8.0.3) 的 Kuhn-Tucker 点. 如果 x_k 是(8.2.1)与(8.2.2)的解, 但(8.2.4)不满足, 我们可令 $i_k\in I_k$, 使得

$$(\lambda_k)_{i_k}=\min_{i\in I_k}(\lambda_k)_i<0, \tag{8.2.5}$$

且取 $S_k:=S_k\setminus\{i_k\}$, 然后重新求解(8.2.1)与(8.2.2). 如果当前迭代点 x_k 不是(8.2.1)与(8.2.2)的稳定点, 我们可求解

$$\min_{d\in\mathbf{R}^n} f(x_k+d) \tag{8.2.6}$$

$$\text{s.t.}\ a_i^Td=0,\ i\in S_k. \tag{8.2.7}$$

设(8.2.6)与(8.2.7)的解为 d_k. 显然 x_k+d_k 是(8.2.1)与(8.2.2)的解. 如果 x_k+d_k 是可行点, 则可令 x_k+d_k 为新的迭代点. 否则, 假定 d_k 是下降方向, 即

$$d_k^Tg_k<0. \tag{8.2.8}$$

我们则可在 d_k 上进行搜索求得 α_k. 只不过是步长 α_k 应满足

$$\alpha_k\leqslant\bar{\alpha}_k=\min_{\substack{i\in I_k\\ a_i^Td_k<0}}\frac{b_i-a_i^Tx_k}{a_i^Td_k}, \tag{8.2.9}$$

才会使 $x_k+\alpha_kd_k$ 是可行点. 值得指出的是: 由于受 $\alpha_k\leqslant\bar{\alpha}_k$ 的限制, 常规的线搜索条件并不一定满足. 例如采用非精确线搜索(2.5.8)与(2.5.9), 但是有可能(2.5.8)与(2.5.9)在$\alpha\in[0,\bar{\alpha}_k]$上无解. 这时我们取 $\alpha_k=\bar{\alpha}_k$.

与二次规划类似, 我们可用线性规划的第一阶段算法求得(8.0.2)与(8.0.3)的一个可行点 x_1 作为初始值, 且令

$$S_1=\mathscr{A}(x_1), \tag{8.2.10}$$

其中 $\mathscr{A}(x_1)$ 由(1.1.12)定义, 是在 x_1处的积极集合. 从而我们可给出求解(8.0.1)与(8.0.2)的积极集法, 算法如下:

算法 8.2.1

步 1　给出可行点 x_1, $S_1:=\mathscr{A}(x_1)$; $k:=1$;

步 2　如果 x_k 不是(8.2.1)与(8.2.2)的解，则转步 4;

步 3　如果(8.2.4)满足，则停，

　　求出 i_k 满足(8.2.5),

$$S_k:=S_k\setminus\{i_k\};$$

步 4　解(8.2.6)与(8.2.7)给出 d_k;

步 5　进行线搜索求出 $\alpha_k>0$ 且满足(8.2.9);

$$x_{k+1}=x_k+\alpha_k d_k,$$

　　对 $j\notin S_k$ 作:

　　　如果 $a_j^T x_{k+1}=b_j$, 则 $S_k:=S_k\cup\{j\}$;

步 6　$S_{k+1}:=S_k$; $k=k+1$; 转步 2.

算法 8.2.1 需要求解(8.2.6)与(8.2.7), 由 8.1 节可知求解(8.2.6)与(8.2.7)需要计算 Z_k. 在不同的迭代中 S_k 不一样，从而导致 $A_k=[a_i](i\in S_k)$ 的不一样. 故每次迭代要重新计算 Z_k, 由于在一般情况下 S_k 和 S_{k+1} 只有一个元素的差别，我们可利用 Z_k 来求得 Z_{k+1}, 同样也可利用已有的近似海色阵或者近似既约海色阵来计算新的近似海色阵或近似既约海色阵.

先考虑加一个约束的情形:

$$S_{k+1}=S_k\cup\{p\}. \tag{8.2.11}$$

我们有 Z_k 使得 $A_k^T Z_k=0$, 秩$(Z_k)=n-|S_k|$, 设有近似

$$\hat{B}_k\approx Z_k^T\nabla^2 f(x_{k+1})Z_k. \tag{8.2.12}$$

我们利用 Z_k 的列向量找到 $n-|S_k|-1$ 个向量和 a_p 正交. 也就是说，令 T 为一非奇异阵，使得

$$Z_k T^{-1}=[Z_{k+1}, u], \tag{8.2.13}$$

$$Z_{k+1}^T a_p=0. \tag{8.2.14}$$

显然 Z_{k+1} 满足 $A_{k+1}^T Z_{k+1}=0$, 秩 $(Z_{k+1})=n-|S_{k+1}|$. 我们将 $\hat{B}_k$ 写成如下形式:

$$\hat{B}_k=Z_k^T B_k Z_k. \tag{8.2.15}$$

则由(8.2.13)可知

$$
\begin{aligned}
T^{-T}\hat{B}_k T^{-1} &= [Z_{k+1} \quad u]^T B_k [Z_{k+1} \quad u] \\
&= \begin{bmatrix} Z_{k+1}^T B_k Z_{k+1} & Z_{k+1}^T B_k u \\ u^T B_k Z_{k+1} & u^T B_k u \end{bmatrix}.
\end{aligned} \tag{8.2.16}
$$

于是我们可取 $\hat{B}_{k+1}$ 为 $T^{-T}\hat{B}_k T^{-1}$ 的前 $n-|S_{k+1}|$ 行和前 $n-|S_{k+1}|$ 列.

记 $\hat{H}_k=\hat{B}_k^{-1}$，考虑 $n\times n$ 矩阵

$$
\overline{H}_k = Z_k \hat{H}_k Z_k^T, \tag{8.2.17}
$$

如果 $Z_k^T Z_k = I$，则有

$$
Z_k^T \overline{H}_k Z_k = (Z_k^T B_k Z_k)^{-1}. \tag{8.2.18}
$$

故可将 $\overline{H}_k$ 理解为 B_k 的逆在 A_k^T 的零空间上的投影，有的算法利用 $\overline{H}_k$ 求解(8.2.1)与(8.2.2)(如 Goldfarb, 1969). 我们可利用 $\overline{H}_k$ 计算 $\overline{H}_{k+1}$:

$$
\begin{aligned}
\overline{H}_{k+1} &= Z_{k+1}\hat{H}_{k+1}Z_{k+1}^T \\
&= [Z_{k+1} \quad u]\begin{bmatrix} \hat{H}_{k+1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} Z_{k+1}^T \\ u^T \end{bmatrix} \\
&= Z_k T^{-1}\begin{bmatrix} \hat{B}_{k+1}^{-1} & 0 \\ 0 & 0 \end{bmatrix} T^{-T} Z_k \\
&= Z_k T^{-1}\left[\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} T^{-T}\hat{B}_k T^{-1}\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}\right]^{+} T^{T-T} Z_k \\
&= \overline{H}_k - \frac{\overline{H}_k a_p a_p^T \overline{H}_k}{a_p^T \overline{H}_k a_p}.
\end{aligned} \tag{8.2.19}
$$

上式的推导依赖于下面的引理:

引理 8.2.2 设对称矩阵 $H\in\mathbf{R}^{(n+1)\times(n+1)}$ 以及它的 n 阶主子式都是可逆的，则有

$$
\left[\begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix} H \begin{pmatrix} I_n & 0 \\ 0 & 0 \end{pmatrix}\right]^{+} = H^{-1} - \frac{H^{-1}e_{n+1}e_{n+1}^T H^{-1}}{e_{n+1}^T H^{-1} e_{n+1}}. \tag{8.2.20}
$$

证明从略.

现在考虑删去一个约束的情形，在这情况下，我们可令

$$
Z_{k+1} = [Z_k \quad \tilde{z}_k]. \tag{8.2.21}
$$

其中

$$\tilde{z}_k = a_{i_k} / \|a_{i_k}\|. \tag{8.2.22}$$

a_{i_k} 是删去的约束的梯度．于是

$$[Z_k \quad \tilde{z}_k]^T B_k [Z_k \quad \tilde{z}_k] = \begin{bmatrix} \hat{B}_k & Z_k^T B_k \tilde{z}_k \\ \tilde{z}_k^T B_k Z_k & \tilde{z}_k^T B_k \tilde{z}_k \end{bmatrix}. \tag{8.2.23}$$

由于设有 $Z_k^T B_k \tilde{z}_k$ 及 $\tilde{z}_k^T B_k \tilde{z}_k$，我们可简单地取

$$\hat{B}_{k+1} = \begin{bmatrix} \hat{B}_k & 0 \\ 0 & 1 \end{bmatrix}. \tag{8.2.24}$$

对于需要计算 $\overline{H}_k$ 的算法，我们可取

$$\begin{aligned} \overline{H}_{k+1} &= Z_{k+1} \begin{bmatrix} \hat{B}_k^{-1} & 0 \\ 0 & 1 \end{bmatrix} Z_{k+1}^T \\ &= Z_k \hat{H}_k Z_k^T + \tilde{z}_k \tilde{z}_k^T \\ &= \overline{H}_k + \frac{a_{i_k} a_{i_k}^T}{\|a_{i_k}\|_2^2}. \end{aligned} \tag{8.2.25}$$

和二次规划的积极集法类似，这里也有可能出现退化的情况．而且处理的方法也类似，所以就不进一步介绍了．

与二次规划的积极集法不一样的地方是问题(8.2.1)与(8.2.2)不像问题(6.3.1)与(6.3.2)那样可经过有限次计算求解．所以在实用的算法中，步 4 的求解 (8.2.6)、(8.2.7) 应换成求(8.2.6)、(8.2.7)的近似解；在步 2 中判断 x_k 是否为 (8.2.1)、(8.2.2)的解应换为判断 x_k 是否为(8.2.1)、(8.2.2)的近似解．

但由于接受(8.2.1)、(8.2.2)近似解，积极集法可能出现锯齿(zigzagging)现象，且收敛于非稳定点的点．关于锯齿现象的例子可见 Wolfe(1972)．克服锯齿现象的方法之一是利用在当前迭代点 x_k 处函数可能下降的估计式：

$$\rho_k = \frac{1}{2} g_k^T Z_k \hat{H}_k Z_k^T g_k, \tag{8.2.26}$$

其中 $\hat{H}_k = \hat{B}_k^{-1}$，$\hat{B}_k$ 是近似既约海色阵．其基本思想是：设 $k' < k$ 是最后一次删去一个约束的迭代下标．如果

$$\rho_k \leqslant f(x_{k'}) - f(x_k). \tag{8.2.27}$$

我们即可把 x_k 当作一个近似解.至于近似求解(8.2.6)与(8.2.7),我们可用

$$\min_{d \in \mathbf{R}^n} g_k^T d + \frac{1}{2} d^T Z_k B_k Z_k^T d, \tag{8.2.28}$$

$$A_k^T d = 0 \tag{8.2.29}$$

来代替(8.2.6)与(8.2.27). 或者可取

$$d_k = -Z_k \hat{H}_k Z_k^T g_k \tag{8.2.30}$$

作为(8.2.6)与(8.2.7)的近似解. 于是算法 8.2.1 的修正算法可给出如下:

算法 8.2.3 (修正积极集法)

步 1 给出可行点 x_1, $S_1 = \mathscr{A}(x_1)$, $\hat{H}_1$ 正定; $\varepsilon \geqslant 0$, $k := 1$; $k' = 1$;

步 2 由(8.2.26)计算 ρ_k,
如果 $\rho_k > f(x_{k'}) - f(x_k)$, 则转步 4,
计算乘子 λ_k, 如果(8.2.4)不满足, 则转步 3,
如果 $\rho_k \leqslant \varepsilon$, 则停,
转步 4;

步 3 求出 i_k 满足(8.2.5),
$S_k := S_k \setminus \{i_k\}$,
$k' := k$;

步 4 求解(8.2.28)与(8.2.29)给出 d_k;

步 5 进行线搜索求出 $\alpha_k > 0$, 且满足(8.2.9);
$x_{k+1} = x_k + \alpha_k d_k$,
利用拟牛顿公式计算 $\hat{B}_{k+1}$,
对 $j \notin S_k$, 作
  如果 $a_j^T x_{k+1} = b_j$ 则 $S_k := S_k \cup \{j\}$;

步 6 $S_{k+1} := S_k$, 计算 Z_{k+1},
$k := k+1$, 转步 2.

在算法的步 2 中, 可能在 x_k 并不是 (8.2.6) 与 (8.2.7)的

Kuhn-Tucker 点时需要计算乘子 λ_k. 这时，我们可取 λ_k 为下面问题

$$\min\|A_k\lambda-g_k\|_2 \tag{8.2.31}$$

的解. 与 Fletcher(1981) 的定理 11.3.1 类似，对于算法 8.2.3 有如下收敛性结果：

定理 8.2.4 设问题 (8.0.1)～(8.0.3) 的可行点集合非空，$f(x)$在可行域上是一致凸函数，如果 $\hat{B}_k$ 正定且$\{\hat{B}_k\}$和$\{\hat{B}_k^{-1}\}$一致有界，而且 A_k 对一切 k 均是列满秩，则对任何可行点 x_1，算法 8.2.3 产生的点列必收敛于问题 (8.0.1)～(8.0.3) 的唯一的解 x^*. 如果严格互补条件在 x^* 满足，则对一切充分大的 k 均有 $S_k=\mathscr{A}(x^*)$，从而不可能出现锯齿现象.

此定理的证明和 Fletcher(1981) 的定理 11.3.1 的证明几乎完全一样，故从略.

§ 8.3 投 影 梯 度 法

本节介绍一个由 Calamai 和 Moré(1987) 给出的解线性约束问题的投影梯度法. 在此记可行集为 X：

$$X=\{x\,|\,a_i^Tx=b_i,\ i\in E;\ a_i^Tx\geqslant b_i,\ i\in I\}. \tag{8.3.1}$$

且定义映照 P 为

$$P(x)=\operatorname{argmin}\{\|z-x\|,\ z\in X\}. \tag{8.3.2}$$

其中 argmin 是指任一 $z\in X$ 使 $\|z-x\|$ 达到最小，$\|\cdot\|$ 是一内积范数. 为了简单起见，我们假定 $\|\cdot\|$ 是欧氏范数 $\|\cdot\|_2$.

投影梯度法也是一个可行点方法. 设当前迭代点为 $x_k\in X$，投影梯度法进行折线搜索，即求 $\alpha_k>0$，使得

$$f(x_k(\alpha_k))\leqslant f(x_k)+\mu_1(x_k(\alpha_k)-x_k)^T\nabla f(x_k), \tag{8.3.3}$$

$$\alpha_k\geqslant\gamma_1 \text{ 或者 } \alpha_k\geqslant\gamma_2\bar{\alpha}_k>0. \tag{8.3.4}$$

其中 $\bar{\alpha}_k>0$ 满足

$$f(x_k(\bar{\alpha}_k))>f(x_k)+\mu_2(x_k(\bar{\alpha}_k)-x_k)^T\nabla f(x_k). \tag{8.3.5}$$

γ_1，γ_2，μ_1，μ_2 是正常数，且 μ_1、$\mu_2\in(0, 1)$. $x_k(\alpha)$是如下折线：

$$x_k(\alpha)=P(x_k-\alpha\nabla f(x_k)). \tag{8.3.6}$$

Calamai 和 Moré 的方法可写成如下形式:

算法 8.3.1

步 1 给出可行点 x_1, $\mu\in(0, 1)$, $\gamma>0$, $\alpha_0=1$, $k:=1$;

步 2 $\alpha_k:=\max\{2\alpha_{k-1}, \gamma\}$,

步 3 如果(8.3.3)式满足,则转步 4,

$\alpha_k:=\alpha_k/4$,

转步 3;

步 4 $x_{k+1}=x_k(\alpha_k)$;

转步 2.

显然, 算法求出的 α_k 满足 (8.3.3)～(8.3.5), 如果我们取 $\mu_2=\mu_1$, $\gamma_1=\gamma$, $\gamma_2=\dfrac{1}{4}$.

由于 $P(x)$ 的定义,对任何 $x\in\mathbf{R}^n$, 有

$$(x-P(x))^T(z-P(x))\leqslant 0, \ \forall z\in X. \tag{8.3.7}$$

于是在(8.3.6)中令 $x=x_k-\alpha_k\nabla f(x_k)$, $z=x_k$,

$$(x_k-\alpha_k\nabla f(x_k)-x_{k+1})^T(x_k-x_{k+1})\leqslant 0. \tag{8.3.8}$$

于是由(8.3.3)和(8.3.8)可知

$$f(x_k)-f(x_{k+1})\geqslant\mu_1\frac{\|x_{k+1}-x_k\|_2^2}{\alpha_k}. \tag{8.3.9}$$

于是有如下结果:

引理 8.3.2 设 $f(x)$在可行域 X 上连续可微且在 X 上下方有界. 如果 $\nabla f(x)$在可行域上一致连续, 则由算法 8.3.1 产生的点列$\{x_k\}$ 必有

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x_k\|}{\alpha_k}=0. \tag{8.3.10}$$

证明 假定引理不真,则存在一个无穷子列 K_0 使得

$$\frac{\|x_{k+1}-x_k\|}{\alpha_k}\geqslant\delta, \quad \forall k\in K_0. \tag{8.3.11}$$

其中 $\delta>0$ 是一个与 k 无关的常数. 由(8.3.9)和(8.3.11)可知对所有 $k\in K_0$ 都有

$$f(x_k)-f(x_{k+1})\geqslant\delta\mu_1\|x_{k+1}-x_k\|\geqslant\delta^2\mu_1\alpha_k. \qquad (8.3.12)$$

因为$f(x)$在可行域上有下界，所以

$$\sum_{k=1}^{\infty}[f(x_k)-f(x_{k+1})]<+\infty. \qquad (8.3.13)$$

利用(8.3.12)和(8.3.13)即知

$$\lim_{\substack{k\to\infty\\ k\in K_0}}\|x_{k+1}-x_k\| = \lim_{\substack{k\in K_0\\ k\to\infty}}\alpha_k=0. \qquad (8.3.14)$$

所以对充分大的$k\in K_0$，(8.3.4)的第一个条件不满足，故有

$$\alpha_k\geqslant\gamma_2\bar{\alpha}_k; \qquad (8.3.15)$$

且(8.3.5)式成立．如果$\bar{\alpha}_k\leqslant\alpha_k$，由于函数

$$\psi(\alpha)=\frac{\|P(x+\alpha d)-x\|}{\alpha},\quad \alpha>0 \qquad (8.3.16)$$

是单调非增的[见 Gafni 和 Bertsekas(1984)或 Calamai 和 Moré (1987)]，我们有

$$\frac{\|x_k-x_k(\bar{\alpha}_k)\|}{\bar{\alpha}_k}\geqslant\frac{\|x_k-x_k(\alpha_k)\|}{\alpha_k}. \qquad (8.3.17)$$

否则，$\bar{\alpha}_k>\alpha_k$，从而$\|x_k-x_k(\bar{\alpha}_k)\|\geqslant\|x_k-x_k(\alpha_k)\|$．故有

$$\frac{\|x_k-x_k(\bar{\alpha}_k)\|}{\bar{\alpha}_k}\geqslant\frac{1}{\gamma_2}\ \frac{\|x_k-x_k(\alpha_k)\|}{\alpha_k}. \qquad (8.3.18)$$

于是在(8.3.5)中令$x=x_k-\bar{\alpha}_k\nabla f(x_k)$，$z=x_k$，可得

$$-(x_k(\bar{\alpha}_k)-x_k)^T\nabla f(x_k)\geqslant\frac{\|x_k-x_k(\bar{\alpha}_k)\|^2}{\bar{\alpha}_k}$$

$$\geqslant\min\left\{1,\ \frac{1}{\gamma_2}\right\}\delta\|x_k-x_k(\bar{\alpha}_k)\|. \qquad (8.3.19)$$

对所有充分大的$k\in K_0$都成立．由于$f(x)$在X上一致连续，我们有

$$f(x_k(\bar{\alpha}_k))-f(x_k)=(x_k(\bar{\alpha}_k)-x_k)^T\nabla f(x_k)+o(\|x_k-x_k(\bar{\alpha}_k)\|). \qquad (8.3.20)$$

由(8.3.20)和(8.3.5)即知

$$-(x_k(\bar{\alpha}_k)-x_k)^T\nabla f(x_k)\leqslant o(\|x_k-x_k(\bar{\alpha}_k)\|). \qquad (8.3.21)$$

(8.3.21)显然与(8.3.19)矛盾．此矛盾说明引理为真．　█

引理 8.3.3　$x^*\in X$ 是问题(8.0.1)～(8.0.3)的 Kuhn-

Tucker点当且仅当存在$\delta>0$，使得

$$P(x^*-\alpha\nabla f(x^*))=x^* \tag{8.3.22}$$

对一切$\alpha\in[0,\ \delta]$均成立.

证明 (8.3.22)等价于

$$\|x^*-\delta\nabla f(x^*)-x^*\|_2^2\leqslant\|x^*-\delta\nabla f(x^*)-x\|_2^2 \tag{8.3.23}$$

对一切$x\in X$均成立. 由于X是凸集，(8.3.23)等价于

$$(x-x^*)^T\nabla f(x^*)\geqslant 0 \tag{8.3.24}$$

对所有充分靠近x^*的可行点x都成立. 因为$(x-x^*)^T\nabla f(x^*)$是x的线性函数，(8.3.24)等价于

$$(x-x^*)^T\nabla f(x^*)\geqslant 0,\ \forall x\in X. \tag{8.3.25}$$

显然(8.3.25)等价于x^*是问题(8.0.1)～(8.0.3)的Kuhn-Tucker点. ▮

利用引理8.3.2和引理8.3.3可得到算法(8.3.1)的收敛性结果.

定理8.3.4 设$f(x)$在可行域X上连续可微，则算法8.3.1产生的点列$\{x_k\}$的任一聚点x^*都是问题(8.0.1)～(8.0.3)的Kuhn-Tucker点.

证明 假定定理不真. 则有$\{x_k\}$的一个收敛子列满足

$$\lim_{\substack{k\in K_0\\k\to\infty}} x_k=x^*, \tag{8.3.26}$$

且

$$P(x^*-\delta\nabla f(x^*))\neq x^*. \tag{8.3.27}$$

其中$\delta>0$，K_0是$\{1,\ 2,\ \cdots\}$的一个子集. 由于(8.3.26)，我们可假定$x_k\in S(k\in K_0)$，其中S是一有界闭集. 因为$\nabla f(x)$在S上一致连续，由引理8.3.2可知

$$\lim_{\substack{k\in K_0\\k\to\infty}}\frac{\|x_{k+1}-x_k\|}{\alpha_k}=0. \tag{8.3.28}$$

由于$\nabla f(x)$的连续性以及(8.3.26)与(8.3.27)，我们有

$$\lim_{\substack{k\in K_0\\k\to\infty}}\frac{\|x_k(\delta)-x_k\|}{\delta}=\frac{\|P(x^*-\delta\nabla f(x^*))-x^*\|}{\delta}>0. \tag{8.3.29}$$

由于(8.3.16)定义的 $\psi(\alpha)$是单调非增的，由(8.3.28)和(8.3.29)可知对所有充分大的 $k\in K_0$ 有

$$\alpha_k \geqslant \bar{\delta}. \tag{8.3.30}$$

故知

$$\begin{aligned} f(x_k)-f(x_{k+1}) &\geqslant -\mu_1(\nabla f(x_k))^T(x_k(\alpha_k)-x_k) \\ &\geqslant -\mu_1(\nabla f(x_k))^T(x_k(\bar{\delta})-x_k) \\ &\geqslant \mu_1 \frac{\|x_k(\bar{\delta})-x_k\|^2}{\bar{\delta}}. \end{aligned} \tag{8.3.31}$$

由(8.3.29)和(8.3.31)即可得到

$$\liminf_{\substack{k\in K_0 \\ k\to\infty}} [f(x_k)-f(x_{k+1})] > 0. \tag{8.3.32}$$

这显然与 $\lim\limits_{k\to\infty} f(x_k)=f(x^*)$相矛盾. 所以定理为真.

§8.4 信 赖 域 法

本节我们给出一个利用信赖域技术的可行点方法. 在第 k 次迭代中设 x_k 是一可行点，我们求解

$$\min_{d\in \mathbb{R}^n} g_k^T d+\frac{1}{2} d_k^T B_k d_k = \varphi_k(d), \tag{8.4.1}$$

$$a_i^T d=0,\ i\in E, \tag{8.4.2}$$

$$a_i^T(x_k+d)\geqslant b_i,\ i\in I, \tag{8.4.3}$$

$$\|d\|_\infty \leqslant \Delta_k. \tag{8.4.4}$$

其中 $\Delta_k>0$ 是信赖域半径. 不难看出，(8.4.1)～(8.4.4)是一个二次规划问题，可用第六章的方法求解. 记 d_k 为(8.4.1)～(8.4.4)的解，下次迭代点 x_{k+1} 以及信赖域半径 Δ_{k+1} 的选取依赖于比值

$$r_k=\frac{f(x_k)-f(x_k+d_k)}{\varphi_k(0)-\varphi_k(d_k)}. \tag{8.4.5}$$

由 d_k 的定义，显然有 $d_k=0$ 当且仅当 x_k 是原问题(8.0.1)～(8.0.3)的 Kuhn-Tucker 点. 由于每次求解子问题把所有的约

束都考虑到了，故不可能像积极集法那样出现锯齿(zigzagging)现象. 下面给出的信赖域法是假定初始点 x_1 是一个可行点.

算法 8.4.1

步 1　给出 x_1 满足(8.0.2)与(8.0.3);

给出 B_1, $\Delta_1>0$, $\varepsilon\geqslant 0$, $k:=1$.

步 2　求解(8.4.1)～(8.4.4)给出 d_k;

如果 $\|d_k\|\leqslant\varepsilon$, 则停;

计算(8.4.5);

$$x_{k+1}=\begin{cases}x_k+d_k, & \text{如果 } r_k>0,\\ x_k, & \text{否则};\end{cases}\tag{8.4.6}$$

步 3　如果 $r_k\geqslant 0.25$, 则转步 4,

$\Delta_k:=\Delta_k/2$, 转步 5;

步 4　如果 $r_k<0.75$ 或者 $\|d_k\|<\Delta_k$, 则转步 5.

$\Delta_k:=2\Delta_k$;

步 5　$\Delta_{k+1}:=\Delta_k$. 计算矩阵 B_{k+1},

$k:=k+1$, 转步 2.

在算法 8.4.1 中，B_{k+1} 可由拟牛顿公式给出. 在下面的收敛性分析中，我们假定$\{B_k\}$一致有界，即存在常数 $M>0$, 使得

$$\|B_k\|\leqslant M.\tag{8.4.7}$$

定理 8.4.2　设 $f(x)$ 在可行域上连续可微，(8.4.7)式满足，算法 8.4.1 产生的点列 $\{x_k\}$ 如有有限聚点，则必有一个聚点是原问题的 Kuhn-Tucker 点.

证明　假定定理不真，我们可证

$$\lim_{k\to\infty}\Delta_k=0.\tag{8.4.8}$$

如果(8.4.8)不成立，则存在常数 $\delta>0$, 使得有无穷多个 k, 使

$$\Delta_k\geqslant\delta \text{ 且 } r_k\geqslant 0.25.\tag{8.4.9}$$

记所有满足(8.4.9)的 k 的集合为 K_0, 不妨设

$$\lim_{\substack{k\in K_0\\ k\to\infty}}x_k=\bar{x}.\tag{8.4.10}$$

根据假设，$\bar{x}$ 不是(8.0.1)～(8.0.3)的 Kuhn-Tucker 点，故 $d=0$

不是

$$\min_{d\in\mathbf{R}^n} g(\bar{x})^T d+\frac{1}{2}M\|d\|_2^2; \tag{8.4.11}$$

$$a_i^T d=0,\ i\in E; \tag{8.4.12}$$

$$a_i^T(\bar{x}+d)\geqslant b_i,\ i\in I; \tag{8.4.13}$$

$$\|d\|_\infty\leqslant\delta \tag{8.4.14}$$

的解. 设 $\bar{d}$ 是(8.4.11)～(8.4.14)的解，则有

$$\eta=g(\bar{x})^T\bar{d}+\frac{1}{2}M\|\bar{d}\|_2^2<0. \tag{8.4.15}$$

于是由(8.4.9)与(8.4.10)以及(8.4.7)可知

$$\varphi_k(0)-\varphi_k(d_k)\geqslant-\frac{1}{2}\eta>0 \tag{8.4.16}$$

对所有充分大的 $k\in K_0$ 均成立，利用(8.4.16)和(8.4.9)的第二个不等式即知

$$f(x_k)-f(x_{k+1})\geqslant-\frac{1}{8}\eta>0 \tag{8.4.17}$$

对所有充分大的 $k\in K_0$ 都成立，由于 $\lim\limits_{k\to\infty}f(x_k)=f(\bar{x})$，(8.4.17)显然是不可能的. 此矛盾说明了只要定理不真则必有(8.4.8).

如果(8.4.8)成立，则必存在一个子列 K_1，使得

$$r_k<0.25,\ \forall k\in K_1. \tag{8.4.18}$$

不妨设

$$\lim_{\substack{k\in K_1\\k\to\infty}} x_k=\hat{x}, \tag{8.4.19}$$

根据假设，$\hat{x}$ 不是(8.0.1)～(8.0.3)的 Kuhn-Tucker 点，令 $\hat{d}$ 是问题

$$\lim_{d\in\mathbf{R}^n} g(\hat{x})^T d+\frac{1}{2}M\|d\|_2^2; \tag{8.4.20}$$

$$a_i^T d=0,\ i\in E; \tag{8.4.21}$$

$$a_i^T(\hat{x}+d)\geqslant b_i,\ i\in I; \tag{8.4.22}$$

$$\|d\|_\infty\leqslant 1 \tag{8.4.23}$$

的解，则必有

$$g(\hat{x})\hat{d}+\frac{1}{2}M\|\hat{d}\|_2^2=\hat{\eta}<0. \tag{8.4.24}$$

于是，由于 $\Delta_k\hat{d}$ 是问题

$$\min_{d\in R^n} g(\hat{x})^T d+\frac{1}{2}M\|d\|_2^2; \tag{8.4.25}$$

$$a_i^T d=0,\ i\in E; \tag{8.4.26}$$

$$a_i^T(\hat{x}+d)\geqslant b_i,\ i\in I; \tag{8.4.27}$$

$$\|d\|_\infty\leqslant \Delta_k \tag{8.4.28}$$

的可行点，则只要 $\Delta_k\leqslant 1$，就有

$$g(\hat{x})^T\hat{d}_k+\frac{1}{2}M\|\hat{d}_k\|_2^2\leqslant \Delta_k\hat{\eta}. \tag{8.4.29}$$

这里 $\hat{d}_k$ 是问题 (8.4.25)～(8.4.28) 的解. 利用 (8.4.19) 和 (8.4.25)～(8.4.29)，即知

$$\varphi_k(0)-\varphi_k(d_k)\geqslant -\frac{1}{2}\hat{\eta}\Delta_k \tag{8.4.30}$$

对所有充分大的 $k\in K_1$ 都成立. 由于 $f(x)$ 连续可微以及 $\{B_k\}$ 一致有界，我们有

$$f(x_k)-f(x_k+d_k)=\varphi_k(0)-\varphi_k(d_k)+o(\|d_k\|). \tag{8.4.31}$$

从(8.4.30)和(8.4.31)可证

$$\lim_{\substack{k\in K_1\\k\to\infty}} r_k=1. \tag{8.4.32}$$

这显然与(8.4.18)矛盾. 此矛盾说明定理为真. ▌

类似无约束优化的信赖域法 [见 Powell (1984 d)]，将条件 (8.4.7) 换成

$$\sum_{k=1}^{\infty}\frac{1}{1+\max\limits_{1\leqslant i\leqslant k}\|B_i\|}=+\infty, \tag{8.4.33}$$

定理 8.4.2 仍成立.

我们定义 $f(x)$ 在 X 上的投影梯度为

$$\nabla_X f(x)=\lim_{\alpha\to 0_+}\frac{P(x-\alpha\nabla f(x))-x}{\alpha}. \tag{8.4.34}$$

不难证明，x^* 是问题(8.0.1)～(8.0.3)的 Kuhn-Tucker 点当且仅当

$$\nabla_X f(x^*)=0. \tag{8.4.35}$$

从定理 8.4.2 的证明可看出，只要 d_k 是(8.4.2)～(8.4.4)的可行点，且满足

$$\varphi_k(0)-\varphi_k(d_k)\geqslant\delta\|\nabla_X f(x_k)\|\min\left\{\Delta_k,\ \frac{1}{\|B_k\|}\right\}, \tag{8.4.36}$$

则算法 8.4.1 仍然有收敛性结果(定理 8.4.2)。

设算法 8.4.1 产生的点列收敛于 x^*，而且在 x^* 处二阶充分条件满足，则只要

$$\lim_{k\to\infty}\|(B_k-\nabla^2 f(x^*))d_k\|/\|d_k\|=0, \tag{8.4.37}$$

则有超线性收敛结果

$$\lim_{k\to\infty}\frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|}=0. \tag{8.4.38}$$

信赖域法的一个优点是它对 B_k 无任何正定假设．如果 $f(x)$ 的二阶导数容易计算．我们可取

$$B_k=\nabla^2 f(x_k). \tag{8.4.39}$$

在这种特殊情况下，与 Fletcher(1980)讨论无约束优化的信赖域法类似，如果算法 8.4.1 产生的点列 x_k 收敛于 x^*，则在 x^* 处二阶必要条件必满足．

Burke、Moré 和 Toraldo(1990)给出了一个基于投影梯度的求解线性约束以及凸约束的信赖域方法．

§8.5 ε-积极集法

积极集的缺点是它根本不考虑在当前迭代点的非积极约束．如果一个非积极集中的约束在当前迭代点的值已经非常小，则很可能会使这次迭代的步长变得非常小．为此我们考虑 ε- 积极集法．ε-积极集法的基本思想是：在积极集法中，把非积极集合中那些约束函数值已经“非常小”的约束看作积极约束．

为了对约束函数值已经“非常小”有精确的定义，我们引入 ε-积极集概念如下：

定义 8.5.1 对于约束优化问题(1.1.1)～(1.1.3)，对任何 $x\in\mathbf{R}^n$ 和 $\varepsilon\geqslant 0$，我们称

$$\mathscr{A}(x,\ \varepsilon)=E\cup I(x,\ \varepsilon) \tag{8.5.1}$$

是在 x 点处的 ε-积极集合，其中 $E=\{1,\ \cdots,\ m_e\}$，

$$I(x,\ \varepsilon)=\{i\,|\,c_i(x)\leqslant\varepsilon,\ i\in I=\{m_e+1,\ \cdots,\ m\}\}. \tag{8.5.2}$$

我们称 $c_i(x)\,(i\in\mathscr{A}(x,\ \varepsilon))$ 是在 x 点的 ε-积极约束.

显然 ε-积极集是积极集的推广，而且对照定义 1.1.3 和定义 8.5.1 可知 0-积极集合就是原始的积极集合. ε-积极集的一个显而易见的性质是以下引理：

引理 8.5.2 对任何 $x\in\mathbf{R}^n$ 和 $\varepsilon_1\geqslant\varepsilon_2\geqslant 0$，有

$$\mathscr{A}(x,\ \varepsilon_2)\subseteq\mathscr{A}(x,\ \varepsilon_1). \tag{8.5.3}$$

Powell(1989) 提出的一个求解问题 (8.0.1)～(8.0.3) 的 ε-积极集法，在第 k 次迭代时求解子问题：

$$\min_{d\in\mathbf{R}^n} g_k^{\mathrm{T}}d+\frac{1}{2}\,d^{\mathrm{T}}B_kd, \tag{8.5.4}$$

$$\text{s.t. } a_i^{\mathrm{T}}d=0,\ i\in E, \tag{8.5.5}$$

$$a_i^{\mathrm{T}}d\geqslant 0,\ i\in I(x_k,\ \varepsilon_k). \tag{8.5.6}$$

其中 x_k 是一个满足约束条件的当前迭代点，$g_k=g(x_k)=\nabla f(x_k)$；$B_k\in\mathbf{R}^{n\times n}$ 是一个近似海色阵，$\varepsilon_k>0$ 是一参数且满足 $\varepsilon_k\to 0$（当 $k\to\infty$ 时）. 记(8.5.4)～(8.5.6)的解为 d_k. 然后在 d_k 方向上进行非精确线性搜索，使条件(2.5.8)与(2.5.9)得到满足. 同§8.2 节中讨论的类似，为了使下一个迭代仍是可行点，我们要求搜索步长满足(8.2.9)式. 在 $\alpha_k=\bar{\alpha}_k$ 时，条件(2.5.8)与(2.5.9)可能不满足.

我们可将 Powell(1989) 的算法整理成下列形式：

算法 8.5.3

步 1 给出可行点 x_1，B_1 正定，$\varepsilon\geqslant 0$，$k:=1$；

步 2 求解(8.5.4)～(8.5.6)，给出 d_k，
　　如果 $\|B_kd_k\|\leqslant\varepsilon$，则转步 4；

步 3 进行线搜索计算 $\alpha_k>0$，

$x_{k+1}=x_k+\alpha_k d_k$,

计算 B_{k+1},

$k:=k+1$, 转步 2;

步 4　如果 $\varepsilon_k\leqslant\varepsilon$, 则停,

$x_{k+1}=x_k$, $k:=k+1$, 转步 2.

在算法中, $\varepsilon_k>0(k=1, 2, \cdots)$是一列趋于 0 的数列. $B_k(k>1)$可由拟牛顿公式计算.

在假定 $f(x)$连续可微且满足 Lipschitz 条件, 即存在常数 L, 使得

$$\|\nabla f(x)-\nabla f(y)\|_2\leqslant L\|x-y\|_2,\ \forall x、y\in X, \tag{8.5.7}$$

以及矩阵 B_k 正定且 $\{B_k\}$、$\{B_k^{-1}\}$ 一致有界的条件下, Powell (1989) 证明了如下收敛性结果:

定理 8.5.4　设 $f(x)$ 在可行域上满足 (8.5.7), 存在常数 $m、M>0$, 使得

$$m\|d\|_2^2\leqslant d^T B_k d\leqslant M\|d\|_2^2 \tag{8.5.8}$$

对一切 k 和一切 $d\in\mathbf{R}^n$ 都成立, 则对任何给定的 $\varepsilon>0$, 算法 8.5.3 必有限终止.

关于定理 8.5.4 的证明以及算法 8.5.3 的一些具体计算的细节, 感兴趣的读者可参阅 Powell(1989).

第 9 章

非线性约束优化

本章我们讨论最一般的非线性规划问题(1.1.1)～(1.1.3)的数值计算方法. 首先我们介绍一个基于变量消去的可行方向法. 许多传统的优化方法, 如 Rosen 的梯度投影法, GRG 方法等都可以看成是特殊形式的可行方向法. 我们还要讨论拉格朗日-牛顿法, 并由此导出逐步二次规划方法. 逐步二次规划方法是解非线性规划问题的一类最重要的方法, 由于篇幅所限, 我们在此介绍两个逐步二次规划方法: 一个是基于 L_1 精确罚函数为价值函数的方法; 另一个是基于 Fletcher 的光滑精确罚函数. 一般说来, 利用光滑罚函数作为价值函数可避免马洛托斯效应(Marotos effect). 我们还介绍既约海色阵方法以及一个信赖域方法. 对于一般约束优化问题, 好的算法仍不多, 有效的通用软件就更奇缺. 由于大多数优化问题都是有约束的, 所以对约束优化计算方法的进一步深入研究是十分必要和非常重要的.

§9.1 可行方向法

可行方向法(feasible direction methods)的基本思想是将第六章和第八章的处理线性约束的广义消去法推广到一般非线性约束.

考虑等式约束问题

$$\min_{x\in\mathbb{R}^n} f(x), \tag{9.1.1}$$

$$\text{s.t.}\quad c(x)=0, \tag{9.1.2}$$

其中 $c(x)=(c_1(x),\cdots,c_m(x))^T$. 设我们有 x 的某一分解:

$$x=\begin{bmatrix}x_B\\x_N\end{bmatrix},\tag{9.1.3}$$

其中 $x_B\in\mathbb{R}^m$, $x_N\in\mathbb{R}^{n-m}$. 于是(9.1.2)可写成

$$c(x_B,\ x_N)=0.\tag{9.1.4}$$

假定我们可从(9.1.4)解出 x_B, 即有

$$x_B=\varphi(x_N).\tag{9.1.5}$$

则(9.1.1)与(9.1.2)等价于求解

$$\min_{x_N\in\mathbb{R}^{n-m}} f(x_B,\ x_N)=f(\varphi(x_N),\ x_N)=\tilde{f}(x_N).\tag{9.1.6}$$

我们称

$$\tilde{g}(x_N)=\nabla_{x_N}\tilde{f}(x_N)\tag{9.1.7}$$

为既约梯度. 不难验证

$$\tilde{g}(x_N)=\frac{\partial}{\partial x_N}f(x_B,\ x_N)+\frac{\partial x_B^T}{\partial x_N}\cdot\frac{\partial}{\partial x_B}f(x_B,\ x_N).\tag{9.1.8}$$

由(9.1.4)可知

$$\frac{\partial x_B^T}{\partial x_N}\cdot\frac{\partial}{\partial x_B}c(x_B,\ x_N)^T+\frac{\partial c(x_B,\ x_N)^T}{\partial x_N}=0.\tag{9.1.9}$$

假定 $\dfrac{\partial c^T}{\partial x_B}$ 非奇异, 则由(9.1.8)与(9.1.9)可得到

$$\tilde{g}(x_N)=\frac{\partial f(x_B,\ x_N)}{\partial x_N}-\frac{\partial c(x_B,\ x_N)^T}{\partial x_N}\left[\frac{\partial c(x_B,\ x_N)}{\partial x_B}\right]^{-1}\frac{\partial f(x_B,\ x_N)}{\partial x_B}.\tag{9.1.10}$$

我们可将既约梯度写成 Lagrange 函数在既约空间上的梯度:

$$\tilde{g}(x_N)=\frac{\partial}{\partial x_N}[f(x)-\lambda^T c(x)],\tag{9.1.11}$$

其中 λ 是满足

$$\frac{\partial f(x)}{\partial x_B}=\frac{\partial c(x)^T}{\partial x_B}\lambda\tag{9.1.12}$$

的乘子.

利用既约梯度, 我们可以构造无约束优化问题(9.1.6)的线搜

索方向，例如可取拟牛顿方向

$$\bar{d}_k = -B_k^{-1}\tilde{g}((x_N)_k). \tag{9.1.13}$$

这里下标 k 表示第 k 次迭代，B_k 是一个近似的既约海色阵，B_1 预先给出，$B_k(k>1)$可由拟牛顿公式(如 BFGS)逐步迭代产生．对于无约束问题(9.1.6)作线搜索等价于对 $f(x)$ 作曲线搜索：

$$\min_{\alpha>0} f(\varphi((x_N)_k+\alpha\bar{d}_k), (x_N)_k+\alpha\bar{d}_k). \tag{9.1.14}$$

所以对给出的试探步长 $\alpha>0$ 需要计算 $\varphi((x_N)_k+\alpha\bar{d}_k)$，这等价于求解

$$c(x_B, (x_N)_k+\alpha_k\bar{d})=0. \tag{9.1.15}$$

我们可用近似牛顿法求解 x_B，即

$$x_B^{(0)}=(x_B)_k, \tag{9.1.16}$$

$$x_B^{(i+1)}=x_B^{(i)}-\left[\frac{\partial c(x_k)^T}{\partial x_B}\right]^{-1}c(x_B^{(i)}, (x_N)_k+\alpha\bar{d}_k). \tag{9.1.17}$$

由于牛顿法是二次收敛的，在一般情况下，用(9.1.17)式迭代几次后就可得到满足误差允许的 x_B．如果经过若干次迭代发现点列 $x_B^{(i)}$ 可能不收敛，则缩小 α，重新进行线搜索．

下面给出可行方向法的计算步骤：

算法 9.1.1

步 1　给出可行点 x_1，$\varepsilon\geqslant0$，$k=1$；

步 2　计算

$$\frac{\partial c^T(x_k)}{\partial x}=\begin{bmatrix}A_B\\A_N\end{bmatrix}, \tag{9.1.18}$$

其中 A_B 非奇异；

由(9.1.12)计算 λ，

由(9.1.11)计算即约梯度 $\tilde{g}_k$；

步 3　如果 $\|\tilde{g}_k\|\leqslant\varepsilon$，则停；

利用某种方式产生搜索方向 $\bar{d}_k$；

步 4　对(9.1.14)进行线搜索给出 $\alpha_k>0$，

$$x_{k+1}=(\varphi_B((x_N)_k+\alpha_k\bar{d}_k), (x_N)_k+\alpha_k\bar{d}_k),$$

$$k:=k+1 \quad 转步 2.$$

和线性约束时的情况类似，不必要把变量分离. 记

$$A_k=A(x_k)=\nabla c(x_k)^T, \tag{9.1.19}$$

设有 QR 分解

$$A_k=[Y_k \quad Z_k]\begin{bmatrix} R_k \\ 0 \end{bmatrix}, \tag{9.1.20}$$

则知对任何 $z_k \in \mathbf{R}^{n-m}$, 有

$$d_k=Z_k z_k \tag{9.1.21}$$

为一线性化可行方向. 设 z_k 为由某种方法(如梯度法、拟牛顿法)产生的搜索方向且满足

$$z_k^T \tilde{g}_k < 0. \tag{9.1.22}$$

其中 $\tilde{g}_k = Z_k^T g_k$ 是既约梯度, 则由(9.1.21)定义的 d_k 是 $f(x)$ 的下降方向. 由于

$$x_k+\alpha d_k \tag{9.1.23}$$

可能不是可行点, 我们用下面的方法将它"可行化":

$$x^{(1)}=x_k+\alpha d_k, \tag{9.1.24}$$

$$x_k^{(i+1)}=x_k^{(i)}-R_k^{-T}c(x_k^{(i)}), \quad i=1, 2, \cdots. \tag{9.1.25}$$

如果我们取

$$z_k=-\tilde{g}_k, \tag{9.1.26}$$

即用最速下降法求既约问题(9.1.6), 这时算法 9.1.1 就是广义既约梯度法(generalized reduced gradient method), 简称 GRG 方法(见 Abadie 和 Carpentier, 1969). 其它的选取 z_k 的方法有利用共轭梯度法、牛顿法以及拟牛顿法(9.1.13)等.

可行方向法可理解为算法沿一线性化可行方向找到一个较好点, 然后将该点利用 Newton-Raphson 方法"可行化"得到一个较好的可行点. 由于每次迭代都要经过若干次 Newton-Raphson 迭代, 可行方向法所需计算量还是相当大的.

§ 9.2 Lagrange-Newton 法

Lagrange-Newton 法是基于 Lagrange 函数的梯度以及约

束函数的 Newton-Raphson 方法. 正由于它用到 Lagrange 函数和 Newton 方法，故被称为 Lagrange-Newton 法.

考虑等式约束问题(9.1.1)与(9.1.2)，由第一章的定义可知：x 是(9.1.1)与(9.1.2) 的 k—T 点当且仅当存在乘子 $\lambda\in\mathbf{R}^m$，使得

$$\nabla f(x)-\nabla c(x)^T\lambda=0, \tag{9.2.1}$$

且 x 是一可行点，即

$$c(x)=0. \tag{9.2.2}$$

联立方程组(9.2.1)与(9.2.2)正好是$(n+m)$个方程$(n+m)$个变量，我们可用 Newton-Raphson 方法求解. 为了使雅可比(Jacobi)矩阵对称，我们将(9.2.1)与(9.2.2)写成如下等价形式：

$$\nabla f(x)-\nabla c(x)^T\lambda=0, \tag{9.2.3}$$

$$-c(x)=0. \tag{9.2.4}$$

解(9.2.3)与(9.2.4)的 Newton-Raphson 法计算增量 δx 和 $\delta\lambda$ 满足：

$$\begin{bmatrix} W(x,\lambda) & -A(x) \\ -A(x)^T & 0 \end{bmatrix}\begin{bmatrix}\delta x\\ \delta\lambda\end{bmatrix}=-\begin{bmatrix}\nabla f(x)-A(x)\lambda \\ -c(x)\end{bmatrix}, \tag{9.2.5}$$

其中

$$A(x)=\nabla c(x)^T, \tag{9.2.6}$$

$$W(x,\lambda)=\nabla^2 f(x)-\sum_{i=1}^{m}\lambda_i\nabla^2 c_i(x). \tag{9.2.7}$$

考虑罚函数

$$P(x,\lambda)=\|\nabla f(x)-\nabla c(x)^T\lambda\|_2^2+\|c(x)\|_2^2, \tag{9.2.8}$$

不难验证，由(9.2.5)定义的 δx 和 $\delta\lambda$ 满足

$$\begin{bmatrix}\delta x\\ \delta\lambda\end{bmatrix}^T\nabla P(x,\lambda)=-2P(x,\lambda)\leqslant 0. \tag{9.2.9}$$

这里 ∇P 是指在(x,λ)空间上的梯度. 下面给出的是利用(9.2.8)进行线搜索的 Lagrange-Newton 法.

算法 9.2.1

步 1　给出 $x_1\in\mathbf{R}^n$, $\lambda_1\in\mathbf{R}^m$, $c_1=0.05$, $\varepsilon\geqslant0$, $k:=1$;

步 2　计算 $P(x_k, \lambda_k)$,

如果 $P(x_k, \lambda_k)\leqslant\varepsilon$, 则停,

在 (x_k, λ_k) 处求解(9.2.5)给出 $(\delta x)_k$、$(\delta\lambda)_k$,

$\alpha_k:=1$;

步 3　如果

$$P(x_k+\alpha_k(\delta x)_k, \lambda_k+\alpha_k(\delta\lambda)_k)\leqslant(1-c_1\alpha_k)P(x_k, \lambda_k), \tag{9.2.10}$$

则转步 4,

$\alpha_k=\alpha_k/4$, 转步 3;

步 4　$x_{k+1}=x_k+\alpha_k(\delta x)_k$,

$\lambda_{k+1}=\lambda_k+\alpha_k(\delta\lambda)_k$,

$k:=k+1$, 转步 2.

由阿米久(Armijo)型线搜索条件(9.2.10)不难证明下面收敛性结果:

定理 9.2.2　设算法 9.2.1 产生的点列 $\{(x_k, \lambda_k)\}$ 有界, 如果 $f(x)$ 和 $c(x)$ 都是二次连续可微, 且矩阵

$$\begin{bmatrix} W(x_k, \lambda_k) & -A(x_k) \\ -A(x_k)^T & 0 \end{bmatrix}^{-1} \tag{9.2.11}$$

一致有界, 则 $\{(x_k, \lambda_k)\}$ 的任何聚点都是方程 $P(x, \lambda)=0$ 的根, 从而 $\{x_k\}$ 的任何聚点都是问题(9.1.1)与(9.1.2)的 K—T 点.

证明　假定 $(\bar{x}, \bar{\lambda})$ 是 $\{(x_k, \lambda_k)\}$ 的一个聚点且

$$P(\bar{x}, \bar{\lambda})>0. \tag{9.2.12}$$

我们设

$$\lim_{\substack{k\in K_0\\ k\to\infty}} x_k=\bar{x}, \quad \lim_{\substack{k\in K_0\\ k\to\infty}} \lambda_k=\bar{\lambda}. \tag{9.2.13}$$

K_0 是 $\{1, 2, \cdots\}$ 的一个无穷子集. 由于

$$P(x_{k+1}, \lambda_{k+1})\leqslant(1-c_1\alpha_k)P(x_k, \lambda_k), \tag{9.2.14}$$

从(9.2.12)式可知

$$\lim_{\substack{k\to\infty\\ k\in K_0}} \alpha_k=0. \tag{9.2.15}$$

根据算法即知

$$P(x_k+\hat{\alpha}_k(\delta x)_k,\quad \lambda_k+\hat{\alpha}_k(\delta\lambda)_k)>(1-c_1\hat{\alpha}_k)P(x_k\lambda_k) \tag{9.2.16}$$

对所有充分大的 $k\in K_0$ 均成立，其中 $\hat{\alpha}_k=4\alpha_k$. 记 $(\overline{\delta x},\ \overline{\delta\lambda})$ 为(9.2.5)在 $(\bar{x},\ \bar{\lambda})$ 处的解，由于(9.2.9)和 $\hat{\alpha}_k\to 0$，我们有

$$\lim_{\substack{k\in K_0\\ k\to\infty}}\frac{P(\bar{x}+\hat{\alpha}_k(\overline{\delta x}),\ -\bar{\lambda}+\hat{\alpha}_k(\delta\lambda))-P(\bar{x},\ \bar{\lambda})}{\hat{\alpha}_k}$$

$$=-2P(\bar{x},\ \bar{\lambda}). \tag{9.2.17}$$

利用 $(x_k,\ \lambda_k)\to(\bar{x},\ \bar{\lambda})$ $(k\in K_0)$ 即知 $((\delta x)_k,\ (\delta\lambda)_k)\to(\overline{\delta x},\ \overline{\delta\lambda})$. 于是对充分大的 $k\in K_0$，有

$$\frac{P(x_k-\hat{\alpha}_k(\delta x)_k,\ \lambda_k+\hat{\alpha}_k(\delta\lambda)_k)-P(x_k,\ \lambda_k)}{\hat{\alpha}_k}$$

$$\leqslant -P(x_k,\ \lambda_k). \tag{9.2.18}$$

由于 $c_1<1$，(9.2.18)显然与(9.2.16)相矛盾. 此矛盾说明

$$P(\bar{x},\ \bar{\lambda})=0 \tag{9.2.19}$$

必成立. 所以定理成立.

关于算法 9.2.1 的收敛速度，我们有如下结果：

定理9.2.3 设算法 9.2.1 产生的点列 $\{x_k\}$ 收敛于 x^*，如果 $f(x)$ 和 $c(x)$ 在 x^* 附近三次连续可微，$A(x^*)$ 是列满秩，而且在 x^* 处二阶充分条件满足，则必有 $\lambda_k\to\lambda^*$，且

$$\left\|\begin{pmatrix}x_{k+1}-x^*\\ \lambda_{k+1}-\lambda^*\end{pmatrix}\right\|=O\left(\left\|\begin{pmatrix}x_k-x^*\\ \lambda_k-\lambda^*\end{pmatrix}\right\|^2\right). \tag{9.2.20}$$

证明 由于算法 9.2.1 实质上是(9.2.3)与(9.2.4)的 Newton-Raphson 方法，于是对充分大的 k，有

$$\left\|\begin{pmatrix}x_k+(\delta x)_k-x^*\\ \lambda_k-(\delta\lambda)_k-\lambda^*\end{pmatrix}\right\|=O\left(\left\|\begin{pmatrix}x_k-x^*\\ \lambda_k-\lambda^*\end{pmatrix}\right\|^2\right). \tag{9.2.21}$$

利用(9.2.21)和 $f(x)$、$c(x)$ 的三次连续可微性，知对所有充分大的 k，(9.2.10)在 $\alpha_k=1$ 时成立. 于是即知(9.2.20)成立.

应当指出的是：(9.2.20)并不等价于常规的二次收敛定义

$$\|x_{k+1}-x^*\|=O(\|x_k-x^*\|^2). \tag{9.2.22}$$

Lagrange-Newton 法的一个重要贡献是在它的基础上发展了逐步二次规划方法，后者已经成为当今求解一般非线性约束优化问题的一类最重要的方法．在下一节中将详细讨论逐步二次规划方法．

首先将(9.2.5)写成如下形式：

$$W(x, \lambda)(\delta x)+\nabla f(x)=A(x)(\lambda+\delta\lambda), \tag{9.2.23}$$

$$c(x)+A(x)^T(\delta x)=0. \tag{9.2.24}$$

从(9.2.23)与(9.2.24)可知 $(\delta x)_k$ 是二次规划问题．

$$\min_{d\in\mathbf{R}^n} d^T\nabla f(x_k)+\frac{1}{2}d^TW(x_k, \lambda_k)d, \tag{9.2.25}$$

$$\text{s. t.}\quad c(x_k)+A(x_k)^Td=0 \tag{9.2.26}$$

的 K—T 点．所以 Lagrange-Newton 法可理解为逐步求解二次规划(9.2.25)与(9.2.26)的方法．λ_1 预先给出，对任何 $k\geqslant1$，有

$$\lambda_{k+1}=\lambda_k+\alpha_k(\bar{\lambda}_k-\lambda_k). \tag{9.2.27}$$

其中 $\bar{\lambda}_k$ 是(9.2.25)与(9.2.26)的 Lagrange 乘子，α_k 是第 k 次迭代的步长．

§9.3　逐步二次规划法

逐步二次规划法(sequential quadratic programming methods)是一类十分重要的求解一般非线性约束优化问题的方法，它简称为 SQP 法．

考虑一般非线性约束问题(1.1.1)～(1.1.3)，类似(9.2.25)与(9.2.26)，我们在第 k 次迭代解子问题

$$\min_{d\in\mathbf{R}^n} g_k^Td+\frac{1}{2}d^TB_kd, \tag{9.3.1}$$

$$\text{s. t.}\quad a_i(x_k)^Td+c_i(x_k)=0,\ i\in E; \tag{9.3.2}$$

$$a_i(x_k)^Td+c_i(x_k)\geqslant0,\ i\in I. \tag{9.3.3}$$

其中　$A(x_k)=[a_1(x_k), \cdots, a_m(x_k)]=\nabla c(x_k)^T$；

$g_k=g(x_k)=\nabla f(x_k)$，　$E=\{1, \cdots, m_e\}$，$I=\{m_e+1, \cdots, m\}$，$B_k\in\mathbf{R}^{n\times n}$ 是 Lagrange 函数的海色阵的近似．记(9.3.1)～(9.3.3)

的解为 d_k. 逐步二次规划法用 d_k 作为线搜索方向. 记 λ_k 为 (9.3.1)～(9.3.3)的 Lagrange 乘子. 故有

$$g_k+B_kd_k=A(x_k)\lambda_k, \tag{9.3.4}$$

$$(\lambda_k)_i\geqslant 0,\ i\in I; \tag{9.3.5}$$

$$\lambda_k^T[c(x_k)+A(x_k)^Td_k]=0. \tag{9.3.6}$$

d_k 的一个很好的性质是：它是许多罚函数的下降方向，例如，对于 L_1 精确罚函数我们有：

引理 9.3.1 设 d_k 是(9.3.1)～(9.3.3)的 K—T 点，λ_k 是相应的 Lagrange 乘子，则对于 L_1 罚函数

$$P(x,\sigma)=f(x)+\sigma\left[\sum_{i=1}^{m_e}|c_i(x)|+\sum_{i=m_e+1}^{m}|c_i(x)_-|\right] \tag{9.3.7}$$

有 $P(x_k+\alpha d_k,\sigma)$ 在 $\alpha=0$ 可微，且

$$P'_\alpha(x_k+\alpha d_k,\sigma)|_{\alpha=0}=-d_k^TB_kd_k-\sigma\|\bar{c}(x_k)\|_1+\lambda_k^Tc(x_k). \tag{9.3.8}$$

其中 $\bar{c}(x)$ 是约束破坏函数，根据(7.1.4)式的定义，如果 $d_k^TB_kd_k>0$ 且 $\sigma\geqslant\|\lambda_k\|_\infty$，则 d_k 是罚函数(9.3.7)在 x_k 处的下降方向.

证明 由于 d_k 满足(9.3.2)与(9.3.3)，显然有

$$P'_\alpha(x_k+\alpha d_k,\sigma)|_{\alpha=0}=g_k^Td_k-\sigma\|\bar{c}(x_k)\|_1. \tag{9.3.9}$$

从(9.3.4)和(9.3.6)可知

$$g_k^Td_k+d_k^TB_kd_k=d_k^TA(x_k)\lambda_k=\lambda_k^Tc(x_k). \tag{9.3.10}$$

利用(9.3.9)与(9.3.10)即知道(9.3.8)成立. 如果 B_k 正定以及 $\sigma\geqslant\|\lambda\|_\infty$ 我们有

$$P'_\alpha(x_k+\alpha d_k,\sigma)|_{\alpha=0}=-d_k^TB_kd_k-\sum_{i=1}^{m}(\sigma-(\lambda_k)_i)|\bar{c}_i(x_k)|<0. \tag{9.3.11}$$

从而引理成立. ∎

下面的算法是 Han(1977)提出的逐步二次规划方法：

算法 9.3.2

步 1 给出 $x_1\in\mathbf{R}^n$, $\sigma>0$, $\delta>0$, $B_1\in\mathbf{R}^{n\times n}$; $\varepsilon\geqslant 1$; $k:=1$;

步 2 求解(9.3.1)与(9.3.2)，给出 d_k;

如果 $\|d_k\|\leqslant\varepsilon$, 则停;

求 $\alpha_k\in[0,\ \delta]$, 使得

$$P(x_k+\alpha_k d_k,\ \sigma)\leqslant\min_{0\leqslant\alpha\leqslant\delta}P(x_k+\alpha d_k,\ \sigma)+\varepsilon_k; \tag{9.3.12}$$

步 3　$x_{k+1}=x_k+\alpha_k d_k$;

计算 B_{k+1};

$k:=k+1$; 转步 2.

(9.3.12)的罚函数是 L_1 精确罚函数, $\varepsilon_k(k=1,\ 2,\ \cdots)$ 是一列非负数列且满足

$$\sum_{k=1}^{\infty}\varepsilon_k<+\infty. \tag{9.3.13}$$

Han(1977)证明了下面收敛性定理:

定理 9.3.3　假定 $f(x)$ 和 $c_i(x)$ 连续可微, 存在常数 $m,\ M>0$, 使得

$$m\|d\|^2\leqslant d^TB_kd\leqslant M\|d\|^2. \tag{9.3.14}$$

对一切 k 和 $d\in\mathbf{R}^n$ 都成立, 如果 $\|\lambda_k\|_\infty\leqslant\sigma$ 对一切 k 均成立, 则算法 9.3.2 产生的点列$\{x_k\}$的任何聚点都是问题(1.1.1)～(1.1.3)的 K—T 点.

证明　假定定理不真, 则存在子列收敛于 $\bar{x}$, 且 $\bar{x}$ 不是 K—T 点. 记

$$\lim_{\substack{k\in K_0\\ k\to\infty}}x_k=\bar{x}. \tag{9.3.15}$$

不失一般性, 可假定

$$\lim_{\substack{k\to K_0\\ k\to\infty}}\lambda_k=\bar{\lambda},\quad \lim_{\substack{k\to K_0\\ k\to\infty}}B_k=\bar{B}. \tag{9.3.16}$$

如果

$$\lim_{\substack{k\in K_0\\ k\to\infty}}\|d_k\|=0, \tag{9.3.17}$$

则由

$$g_k+B_kd_k=A_k\lambda_k \tag{9.3.18}$$

可推出

$$g(\bar{x}) = A(\bar{x})\bar{\lambda}. \tag{9.3.19}$$

这与 $\bar{x}$ 不是 K—T 点相矛盾. 故我们可假设

$$\|d_k\| \geqslant \eta > 0, \ \forall k \in K_0. \tag{9.3.20}$$

其中 η 是一常数. 利用(9.3.20)、(9.3.8)以及 $\|\lambda_k\|_\infty \leqslant \sigma$ 可知

$$P'_\alpha(x_k + \alpha d_k, \ \sigma)|_{\alpha=0} \leqslant -m\eta \|d_k\| \tag{9.3.21}$$

对一切 $k \in K_0$ 都成立. 利用(9.3.21)以及函数 $f(x)$ 和 $c(x)$ 的连续可微性可知必存在正常数 $\bar{\eta}$, 使得

$$\min_{0<\alpha<\delta} P(x_k + \alpha d_k, \ \sigma) \leqslant P(x_k, \ \sigma) - \bar{\eta} \tag{9.3.22}$$

对一切 $k \in K_0$ 都成立. 故知

$$P(x_{k+1}, \ \sigma) \leqslant P(x_k, \ \sigma) - \bar{\eta} + \varepsilon_k, \ \forall k \in K_0. \tag{9.3.23}$$

于是我们得到

$$\begin{aligned} \sum_{k \in K_0} \bar{\eta} &\leqslant \sum_{k \in K_0} [P(x_k, \ \sigma) - P(x_{k+1}, \ \sigma)] + \sum_{k \in K_0} \varepsilon_k \\ &\leqslant \sum_{k=1}^{\infty} [P(x_k, \ \sigma) - P(x_{k+1}, \ \sigma)] + \sum_{k=1}^{\infty} \varepsilon_k. \end{aligned} \tag{9.3.24}$$

由于 $\lim\limits_{k \to +\infty} P(x_k, \ \sigma) = P(\bar{x}, \ \sigma)$, 故知不等式(9.3.24)右端的第一个级数收敛. 由(9.3.13)知 $\sum\limits_{k=1}^{\infty} \varepsilon_k$ 有界. 于是从(9.3.24)和 $\bar{\eta} > 0$ 知 K_0 是一个有限集合, 这与假定 K_0 是无穷子列相矛盾. 此矛盾说明定理为真. ∎

利用二次规划子问题的解 d_k 作为搜索方向的好处是: d_k 在一定条件是一个超线性收敛步. 事实上, 对于等式约束问题, Boggs、Tolle 和 Wang(1982)证明了在一定条件下

$$x_k + d_k - x^* = o(\|x_k - x^*\|) \tag{9.3.25}$$

等价于

$$\frac{\|P_k(B_k - W(x^*, \ \lambda^*))d_k\|}{\|d_k\|} \to 0, \tag{9.3.26}$$

其中 $W(x^*, \lambda^*)$ 是由(9.2.7)定义的. P_k 是从 $\mathbf{R}^n$ 投影到 $A(x_k)$ 的零空间上的投影算子:

$$P_k = I - A(x_k)[A(x_k)^T A(x_k)]^{-1} A(x_k)^T. \tag{9.3.27}$$

Powell (1978a)对每个约束函数 $c_i(x)$ 用不同的罚因子而且

提出了一套每次迭代修改这些罚因子的方法. 但是 Powell 的修正技巧并不能满足定理 9.3.3 的条件. Chamberlain (1979) 举例说明 Powell 的修正罚因子的技巧可能导致死循环.

在算法 9.3.2 中, 线搜索是基于 L_1 精确罚函数. 在约束优化中, 线搜索中用到的函数是用来判断试探点好坏的, 所以这个函数也被称为价值函数(merit function). 用 L_1 罚函数作为价值函数的一个困难是有名的马洛托斯效应(Marotos effect), 就是说 L_1 罚函数有可能破坏超线性收敛(Marotos, 1978). 马洛托斯效应可由下例说明:

$$\min \quad 3v^2-2u, \tag{9.3.28}$$

$$\text{s. t.} \quad u-v^2=0. \tag{9.3.29}$$

显然原点$(0, 0)^T$是唯一的极小点且二阶充分条件满足. 对任何靠近解的点($\varepsilon>0$ 充分小)

$$\bar{x}(\varepsilon)=(u(\varepsilon),\ v(\varepsilon))^T=(\varepsilon^2,\ \varepsilon)^T \tag{9.3.30}$$

取 $B=W(x^*, \lambda^*)$, 求解二次规划后可得到步长

$$\bar{d}(\varepsilon)=(-2\varepsilon^2,\ -\varepsilon)^T. \tag{9.3.31}$$

所以我们有

$$\|\bar{x}(\varepsilon)+\bar{d}(\varepsilon)-x^*\|=O(\|\bar{x}(\varepsilon)-x^*\|^2). \tag{9.3.32}$$

因此 $\bar{d}(\varepsilon)$ 是一超线性收敛步. 但是

$$f(\bar{x}(\varepsilon)+\bar{d}(\varepsilon))=2\varepsilon^2, \tag{9.3.33}$$

$$c(\bar{x}(\varepsilon)+\bar{d}(\varepsilon))=-\varepsilon^2. \tag{9.3.34}$$

由于

$$f(\bar{x}(\varepsilon))=\varepsilon^2, \tag{9.3.35}$$

$$c(\bar{x}(\varepsilon))=0. \tag{9.3.36}$$

所以对任何罚因子 $\sigma>0$ 都有

$$P(\bar{x}(\varepsilon)+\bar{d}(\varepsilon),\ \sigma)>P(\bar{x}(\varepsilon),\ \sigma). \tag{9.3.37}$$

故利用 L_1 罚函数作为价值函数, 则 $\bar{x}(\varepsilon)+\bar{d}(\varepsilon)$不能被接受.

近年来, 围绕如何克服马洛托斯效应进行了许多工作. 目前大家已知马洛托斯效应是由于价值函数(9.3.7)的非光滑性引起的.

克服马洛托斯效应的一种办法是引进一个二阶校正步．二阶校正步可以是下述问题的解：

$$\min_{d\in\mathbf{R}^n} \hat{g}_k^T d + \frac{1}{2} d^T B_k d, \tag{9.3.38}$$

$$\text{s.t.} \quad c_i(x_k) + a_i(x_k)^T d = 0, \ i\in E, \tag{9.3.39}$$

$$c_i(x_k) + a_i(x_k)^T d \geqslant 0, \ i\in I. \tag{9.3.40}$$

其中

$$\hat{g}_k = g_k + \frac{1}{2}\sum_{i=1}^{m} (\lambda_k)_i [\nabla c_i(x_k) - \nabla c_i(x_k + d_k)]. \tag{9.3.41}$$

λ_k 是二次规划(9.3.1)～(9.3.3)的 Lagrange 乘子．记(9.3.38)～(9.3.40)的解为 $\hat{d}_k$．在一定条件下可以证明下列超线性收敛结果：

$$\| x_k + d_k + \hat{d}_k - x^* \| / \| x_k - x^* \| \to 0. \tag{9.3.42}$$

且当 x_k 充分靠近 x^* 时，新点 $x_k + d_k + \hat{d}_k$ 可以被价值函数(9.3.7)所接受．关于二阶校正步方法的详细讨论可参阅 Mayne 和 Polak(1982)以及 Fukushima(1986)．这种方法在本质上和在下一章介绍的非光滑优化的二阶校正步方法是一样的．(9.3.38)～(9.3.40)的缺点是需要计算在点 $x_k + d_k$（非迭代点）上约束函数的梯度．这个现象和简约海色阵法及用两个搜索方向的方法中的两步超线性收敛结果有些相似，这是不可避免的，因为不知解的任何信息，要想得到一个可以被 L_1 价值函数所接受的二阶校正步，必须在试探点 $x_k + d_k$ 上知道一些东西．基于下一章非光滑优化的二阶校正步方法，我们建议用下面问题来求解一个二阶校正步：

$$\min \quad g_k^T d + \frac{1}{2} d^T B_k d, \tag{9.3.43}$$

$$\text{s.t.} \quad c_i(x_k + d_k) + a_i(x_k)^T (d - d_k) = 0, \ i\in E, \tag{9.3.44}$$

$$c_i(x_k + d_k) + a_i(x_k)^T (d - d_k) \geqslant 0, \ i\in I. \tag{9.3.45}$$

子问题(9.3.43)～(9.3.45)和(9.3.1)～(9.3.3)的差别只在于线

性约束的常数项变了．而且(9.3.43)～(9.3.45)不像(9.3.38)～(9.3.40)那样要求计算 $\nabla c(x_k+d_k)$．

另一种保持 SQP 方法的超线性收敛的思想是修改接受试探步的条件．粗略地说，既然试探步 d_k 是一超线性收敛步．我们应在保证收敛的前提下尽可能地接受 $\alpha_k=1$ 的步长因子．

Chamberlain 等人(1982)提出 Watchdog 技术就是在一些迭代中利用拉格朗日函数

$$L(x, \lambda)=f(x)-\sum_{i=1}^{m}\lambda_i c_i(x) \tag{9.3.46}$$

作为价值函数，从而放宽了选取 $\alpha_k=1$ 的条件．由于 Watchdog 技术在总体上还是利用 L_1 罚函数判别点的好坏，所以它的总体收敛性仍可保证．因为它在一些迭代放宽线搜索条件，所以在一定条件下可证明它是超线性收敛的．

修改算法 9.3.2 接受试探步条件的另一方法是用其他性质好的罚函数作为价值函数．考虑等式约束问题(9.1.1)与(9.1.2)，Powell 和 Yuan (1986)用 Fletcher 光滑精确罚函数(7.2.38)作为价值函数．由于函数(7.2.38)的导数需要计算 $f(x)$ 和 $c(x)$ 的二阶导数，Powell 和 Yuan 利用(7.2.38)的一个逼近形式

$$\begin{aligned}\Phi_{k,i}(\alpha\beta_{k,i})=&f(x_k+\alpha\beta_{k,i}d_k)-[\lambda(x_k)\\&+\alpha(\lambda(x_k+\beta_{k,i}d_k)-\lambda(x_k))]^T c(x_k+\alpha\beta_{k,i}d_k)\\&+\frac{1}{2}\sigma_{k,i}\|c(x_k+\alpha\beta_{k,i}d_k)\|_2^2,\ 0\leqslant\alpha\leqslant 1.\end{aligned} \tag{9.3.47}$$

其中 d_k 是二次规划(9.3.1)～(9.3.3)的解，$\beta_{k,i}$ 是第 k 次迭代中的第 $i+1$ 个试探步长．$\sigma_{k,i}$ 是当前的罚因子，且满足

$$\begin{aligned}\Phi'_{k,i}(0)&\leqslant-\frac{1}{2}[d_k^T B_k d_k+\sigma_{k,i}\|c(x_k)\|_2^2]\\&\leqslant-\frac{1}{4}\sigma_{k,i}\|c(x_k)\|_2^2.\end{aligned} \tag{9.3.48}$$

于是我们可给出 Powell 和 Yuan(1986)的方法如下：

算法 9.3.4

步 1　给出 $x_1\in\mathbf{R}^n$, $\beta_1\in(0,1)$, $\beta_2\in[\beta_1,1)$,
$\mu\in(0,0.5)$, $\sigma_{1,-1}>0$, $B_1\in\mathbf{R}^{n\times n}$, $\varepsilon\geqslant 0$.
$k:=1$.

步 2　解(9.3.1)～(9.3.3)，给出 d_k;
如果 $\|d_k\|\leqslant\varepsilon$，则停；
令 $i=0$, $\beta_{k,0}=1$;

步 3　选取 $\sigma_{k,i}$ 使得(9.3.48)成立；
如果

$$\phi_{k,i}(\beta_{k,i})\leqslant\phi_{k,i}(0)+\mu\beta_{k,i}\Phi'_{k,i}(0),\tag{9.3.49}$$

则转步 4;
$i:=i+1$; $\beta_{k,i}\in[\beta_1,\beta_2]\beta_{k,i-1}$; 转步 3;

步 4　$x_{k+1}=x_k+\beta_{k,i}d_k$;
$\sigma_{k+1,-1}=\sigma_{k,i}$;
计算 B_{k+1};
$k:=k+1$; 转步 2.

引理 9.3.5　设 $\{x_k\}$、$\{d_k\}$、$\{B_k\}$ 有界，$A(x)=\nabla c(x)^T$ 对一切 $x\in\mathbf{R}^n$ 均为列满秩以及存在常数 $\delta>0$，使得对一切 k 都有

$$d^TB_kd\geqslant\delta\|d\|_2^2,\ \forall A(x_k)^Td=0.\tag{9.3.50}$$

则存在 k' 使得对一切 $k\geqslant k'$ 均有

$$\sigma_{k,i}=\sigma_{k',0}=\bar{\sigma}>0,\tag{9.3.51}$$

而且

$$\|d_k\|\to 0.\tag{9.3.52}$$

利用引理 9.3.5 即可证以下定理：

定理 9.3.6　在引理 9.3.5 的条件下，算法 9.3.4 产生的点列 $\{x_k\}$ 的任何聚点都是问题(9.1.1)与(9.1.2)的 K—T 点.

引理 9.3.5 和定理 9.3.6 的证明可见 Powell 和 Yuan (1986). 利用可微精确罚函数作为价值函数的好处是：靠近解时任何超线性试探步都能被接受.

引理 9.3.7　设引理 9.3.5 的条件满足，算法 9.3.4 产生的点列 $\{x_k\}$ 收敛于 x^*. 设任一子列 $\{k_i, i=1,2,\cdots\}$ 当 $k_i\to\infty$ 时有

$$\|x_{k_i}+d_{k_i}-x^*\|=o(\|x_{k_i}-x^*\|), \tag{9.3.53}$$

则对所有充分大的 i, 有

$$x_{k_i+1}=x_{k_i}+d_{k_i}. \tag{9.3.54}$$

证明　不失一般性，我们可假定所有 $k_i\geqslant k'$. 为了记号简便起见，我们用 t 代替 k_i. 显然只需证明

$$\Phi_{t,0}(1)-\Phi_{t,0}(0)-\mu\Phi'_{t,0}(0)<0. \tag{9.3.55}$$

由(9.3.51)知

$$\Phi_{t,0}(1)=f(x_t+d_t)-\lambda(x_t+d_t)^Tc(x_t+d_t)$$
$$+\frac{1}{2}\bar{\sigma}\|c(x_t+d_t)\|_2^2. \tag{9.3.56}$$

利用 $f(x)$ 的二次连续可微性，我们有

$$f(x_t+d_t)=f(x_t)+\frac{1}{2}d_t^T[g_t+g(x_t+d_t)]+o(\|d_t\|_2^2)$$
$$=f(x_t)+\frac{1}{2}d_t^T[g_t+g(x^*)]+o(\|d_t\|_2^2). \tag{9.3.57}$$

同样，对 $c_i(x_t+d_t)$ 也有类似(9.3.57)的表达式，将这些展开式代入(9.3.56)式中得到

$$\begin{aligned}
&\Phi_{t,0}(1)-\Phi_{t,0}(0)\\
&=\frac{1}{2}d_t^T[g_t+g(x^*)]\\
&\quad-\lambda(x_t+d_t)^T\left[c_t+\frac{1}{2}A_t^Td_t+\frac{1}{2}A(x^*)^Td_t\right]\\
&\quad-\left[-\lambda_t^Tc_t+\frac{1}{2}\bar{\sigma}\|c_t\|_2^2\right]+o(\|d_t\|_2^2)\\
&=\frac{1}{2}\Phi'_{t,0}(0)+\frac{1}{2}d_t^T[g(x^*)-A(x^*)\lambda(x_t+d_t)]\\
&\quad+o(\|d_t\|_2^2)\\
&=\frac{1}{2}\Phi'_{t,0}(0)+o(\|d_t\|_2^2).
\end{aligned} \tag{9.3.58}$$

不难证明存在正常数 $\bar{\eta}>0$, 使得对一切 k 和 i 都有

$$\Phi'_{k,i}(0)\leqslant-\bar{\eta}\|d_k\|^2. \tag{9.3.59}$$

利用(9.3.58)、(9.3.59)以及 $\mu<\frac{1}{2}$ 即知，对充分大的 t (9.3.55)成立. ∎

从引理 9.3.7 可直接推出以下定理：

定理 9.3.8 设引理 9.3.5 的条件满足. 算法 9.3.4 产生的点列$\{x_k\}$收敛于 x^*，如果

$$\lim_{k\to\infty}\frac{\|x_k+d_k-x^*\|}{\|x_k-x^*\|}=0, \tag{9.3.60}$$

则对充分大的 k 都有 $x_{k+1}=x_k+d_k$. 于是点列$\{x_k\}$超线性收敛于 x^*.

其他关于利用可微精确罚函数的方法可参阅 Bertsekas (1982a)、Boggs 和 Wang(1985)、Di Pillo 和 Grippo(1986)等.

另外，利用增广拉格朗日函数作为价值函数的逐步二次规划方法由 Gill、Murray、Saunders 和 Wright(1983) 以及 Schittkowski(1981, 1983)给出.

下面讨论逐步二次规划法中 B_{k+1} 的计算. 一般是用拟牛顿法修正. 由于希望 B_{k+1} 是 Lagrange 函数的海色阵的近似，我们可取

$$s_k=x_{k+1}-x_k, \tag{9.3.61}$$

$$y_k=\nabla f(x_{k+1})-\nabla f(x_k)-\sum_{i=1}^{m}(\lambda_k)_i[\nabla c_i(x_{k+1})-\nabla c_i(x_k)]. \tag{9.3.62}$$

然后利用第四章讨论的拟牛顿公式计算 B_{k+1}. 与无约束优化本质不一样的是：对于价值函数进行线搜索并不能保证

$$s_k^T y_k>0. \tag{9.3.63}$$

从而不能直接利用 BFGS 方法. Powell(1978a, b)建议取

$$\bar{y}_k=\begin{cases} y_k, & \text{如果 } s_k^T y_k\geqslant 0.2s_k^T B_k s_k; \\ \theta_k y_k+(1-\theta_k)B_k s_k, & \text{否则}, \end{cases} \tag{9.3.64}$$

其中

$$\theta_k=\frac{0.8s_k^T B_k s_k}{s_k^T B_k s_k-s_k^T y_k}. \tag{9.3.65}$$

这种选取 $\bar{y}_k$ 的基本思想是利用 y_k 和 B_ks_k 的凸组合构造一可以用来修正矩阵的向量. 由于 B_ks_k 可理解为 y_k 的一种近似估计，且满足(假定 B_k 正定)

$$s_k^T(B_ks_k)>0. \quad (9.3.66)$$

故利用 y_k 和 B_ks_k 的凸组合是一种很自然的选择. Powell 的公式(9.3.64)在几何上可理解为：设 B_ks_k 在 s_k 方向上投影长度为 1，修正公式(9.3.64)实际上就是要求 $\bar{y}_k$ 是在 y_k 和 B_ks_k 的连线上使其尽可能靠近 y_k，且在 s_k 方向上投影的长度至少为 0.2. 这可由图 9.3.1 表明.

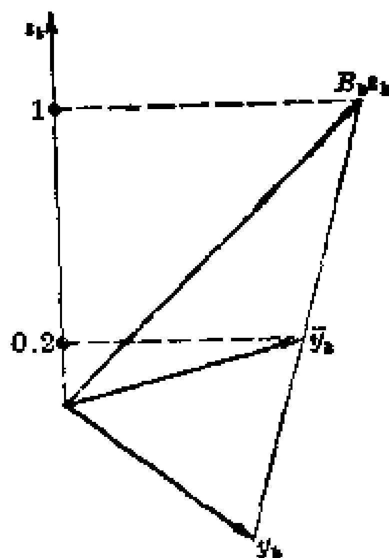


图 9.3.1

得到了修正方向 $\bar{y}_k$ 后，我们可用 BFGS 方法计算 B_{k+1}:

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{\bar{y}_k\bar{y}_k^T}{s_k^T\bar{y}_k}. \quad (9.3.67)$$

另一种修正 y_k 的方法是取

$$\hat{y}_k=y_k+2\sigma\sum_{i=1}^{m}-c_i(x_k)\nabla c_i(x_k) \quad (9.3.68)$$

代替 y_k. 由于

$$\hat{y}_k\approx[\nabla^2L(x_k, \lambda_k)+2\sigma A_kA_k^T]s_k, \quad (9.3.69)$$

所以可理解为利用 $\hat{y}_k$ 实质上是希望 B_{k+1} 近似增广拉格朗日函数的海色阵. 这一方法的优点是

$$s_k^T\hat{y}_k>0, \quad (9.3.70)$$

通常可以得到满足. 如果 $s_k^T\hat{y}_k\leqslant 0$，也可利用增大 σ 的值来使(9.3.70)成立. 而且一般说来，在解处增广拉格朗日函数的海色阵是正定的，故用正定矩阵 B_k 近似它是比较合理的.

§9.4 既约海色阵方法

既约海色阵方法是利用既约海色阵或者是近似既约海色阵构

造的求解约束优化的方法.

考虑等式约束优化问题(9.1.1)与(9.1.2), 假定$f(x)$和$c_i(x)$的二阶导数可计算,则从9.1节知Lagrange-Newton法是一个利用Lagrange函数海色阵的方法. 记

$$W_k=W(x_k, \lambda_k)=\nabla^2_{xx}L(x_k, \lambda_k), \tag{9.4.1}$$

$$A_k=A(x_k)=\nabla c(x_k)^T, \tag{9.4.2}$$

$$g_k=\nabla f(x_k)=\nabla f(x_k). \tag{9.4.3}$$

则Lagrange-Newton法的试探步为$(d_k, (\delta\lambda)_k)$

$$\begin{bmatrix} W_k & -A_k \\ -A_k^T & 0 \end{bmatrix}\begin{bmatrix} d_k \\ (\delta\lambda)_k \end{bmatrix}=-\begin{bmatrix} g_k-A_k\lambda_k \\ -c_k \end{bmatrix}. \tag{9.4.4}$$

令$\hat{\lambda}_k=\lambda_k+(\delta\lambda)_k$,则(9.4.4)可写成

$$\begin{bmatrix} W_k & -A_k \\ -A_k^T & 0 \end{bmatrix}\begin{bmatrix} d_k \\ \hat{\lambda}_k \end{bmatrix}=\begin{bmatrix} -g_k \\ c_k \end{bmatrix} \tag{9.4.5}$$

设A_k的QR分解有如下形式:

$$A_k=[Y_k \quad Z_k]\begin{bmatrix} R_k \\ 0 \end{bmatrix} \tag{9.4.6}$$

利用(9.4.6),我们可将(9.4.5)写成等价形式:

$$\begin{bmatrix} Y_k^TW_kY_k & Y_k^TW_kZ_k & -R_k \\ Z_k^TW_kY_k & Z_k^TW_kZ_k & 0 \\ -R_k^T & 0 & 0 \end{bmatrix}\begin{bmatrix} Y_k^Td_k \\ Z_k^Td_k \\ \hat{\lambda}_k \end{bmatrix}=\begin{bmatrix} -Y_k^Tg_k \\ -Z_k^Tg_k \\ c_k \end{bmatrix}. \tag{9.4.7}$$

令

$$p_k=Y_k^Td_k, \tag{9.4.8}$$

$$q_k=Z_k^Td_k. \tag{9.4.9}$$

显然,p_k和q_k分别是搜索方向d_k在秩空间和零空间上的投影,由于线性方程组(9.4.7)的分块三角形状, 我们很容易依次求解p_k、q_k和$\hat{\lambda}_k$:

$$R_k^Tp_k=Y_k^Tg_k, \tag{9.4.10}$$

$$(Z_k^TW_kZ_k)q_k=-Z_k^Tg_k-Z_k^TW_kY_kp_k, \tag{9.4.11}$$

$$R_k\hat{\lambda}_k=Y_k^Tg_k+Y_k^TW_k(Y_kp_k+Z_kq_k). \tag{9.4.12}$$

如果把(9.4.7)式的后两行(分块意义下的)单独考虑，则得到一个与 λ 无关的线性方程组：

$$\begin{bmatrix} Z_k^T W_k Y_k & Z_k^T W_k Z_k \\ -R_k^T & 0 \end{bmatrix} \begin{bmatrix} p_k \\ q_k \end{bmatrix} = \begin{bmatrix} -Z_k^T g_k \\ c_k \end{bmatrix}. \quad (9.4.13)$$

实质上就是

$$\begin{bmatrix} Z_k^T W_k & Z_k^T W_k \\ -A_k^T & 0 \end{bmatrix} d_k = \begin{bmatrix} -Z_k^T g_k \\ c_k \end{bmatrix}. \quad (9.4.14)$$

Nocedal 和 Overton(1985) 建议用拟牛顿修正矩阵 B_k 代替

$$Z_k^T W_k, \quad (9.4.15)$$

即每次迭代的搜索方向 d_k 是

$$\begin{bmatrix} B_k \\ -A_k^T \end{bmatrix} d = \begin{bmatrix} -Z_k^T g_k \\ c_k \end{bmatrix} \quad (9.4.16)$$

的解，$B_k \in \mathbf{R}^{(n-m)\times n}$ 是 $Z_k^T W_k$ 的一个近似．由于 $Z_k^T W_k$ 是单边既约海色阵(one sided reduced Hessian)，所以这方法也称为单边既约海色阵方法.我们可以用 Broyden 非对称秩为 1 的公式修正 B_k，即

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T}{s_k^T s_k}. \quad (9.4.17)$$

其中

$$s_k = x_{k+1} - x_k, \quad (9.4.18)$$

$$y_k = Z_{k+1}^T g_{k+1} - Z_k^T g_k. \quad (9.4.19)$$

在一定条件下，Nocedal 和 Overton(1985)证明了利用修正公式(9.4.17)～(9.4.19)的单边既约海色阵方法是局部超线性收敛的．

现在考虑双边既约海色阵 (two sided reduced Hessian) 方法．考虑用一个近似既约海色阵 B_k 代替 $Z_k^T W_k Z_k$，而用零矩阵代替 $Z_k^T W_k Y_k$，于是从(9.4.13)式得到

$$\begin{bmatrix} 0 & B_k \\ -B_k^T & 0 \end{bmatrix} \begin{bmatrix} p_k \\ q_k \end{bmatrix} = \begin{bmatrix} -Z_k^T g_k \\ c_k \end{bmatrix}. \quad (9.4.20)$$

这样作的理由是：首先，Powell(1978b)在研究逐步二次规划方法的超线性收敛性质时发现 $Y_k^T W_k Z_k$ 这一矩阵用任一有界矩阵代

替都有两步超线性收敛结果:

$$\frac{\|x_{k+1}-x^*\|}{\|x_{k-1}-x^*\|}\to 0. \tag{9.4.21}$$

另一个理由是当 $c_i(x)$ 是线性函数时，只要 x_k 是可行点，则必有 $p_k=0$ 从而 $Z_k^TW_kY_k$ 并不影响 q_k 的值(见(9.4.13)式)．还有由于 $Z_k^TW_kZ_k$ 是方阵，在解处一般是正定的，所以我们可用正定矩阵(如 BFGS 修正公式)来逼近它.

我们可将(9.4.20)写成等价形式:

$$\begin{bmatrix} B_kZ_k^T \\ -A_k^T \end{bmatrix} d_k=\begin{bmatrix} -Z_k^Tg_k \\ c_k \end{bmatrix}, \tag{9.4.22}$$

B_{k+1} 可由 BFGS 公式修正:

$$B_{k+1}=B_k-\frac{B_ks_ks_k^TB_k}{s_k^TB_ks_k}+\frac{y_ky_k^T}{s_k^Ty_k}. \tag{9.4.23}$$

其中

$$s_k=Z_k(x_{k+1}-x_k), \tag{9.4.24}$$

$$y_k=Z_{k+1}g_{k+1}-Z_kg_k. \tag{9.4.25}$$

双边既约海色阵方法的收敛速度是两步超线性收敛．事实上，有以下定理[见 Coleman 和 Conn (1984); Nocedal 和 Overton(1985)以及 Powell(1978b)]:

定理 9.4.1　设 d_k 由(9.4.22)的定义，$x_{k+1}=x_k+d_k$，$x_k\to x^*$，$\|B_k^{-1}\|$ 一致有界，且在 x^* 处二阶充分条件满足，则 x_k 两步 Q-超线性收敛于 x^*:

$$\|x_{k+1}-x^*\|/\|x_{k-1}-x^*\|\to 0 \tag{9.4.26}$$

当且仅当

$$\frac{\|(B_k-Z(x^*)^TW(x^*,\lambda^*)Z(x^*))Z_k^Td_k\|}{\|d_k\|}\to 0. \tag{9.4.27}$$

而且两步超线性收敛的结果不能进一步改进，Yuan(1985c)利用函数

$$f(x)=\frac{1}{2}z^2-yz+\frac{1}{6(1-z)^3}\Big[-4(z-y)^3-6(z-y)^2(y-z^2)$$

$$-12(z-y)(y-z^2)^2-17(y-z^2)^3+3\frac{(y-z^2)^4}{1-z}\Big], \tag{9.4.28}$$

$$c(x)=y+\frac{1}{(1-z)^3}[(z-y)^2+(z-y)(y-z^2)+2(y-z^2)^2]. \tag{9.4.29}$$

其中 $x=(y,\ z)^T\in\mathbf{R}^2$，构造出

$$B_k=Z(x^*)^TW(x^*,\ \lambda^*)Z(x^*), \tag{9.4.30}$$

得到点列$\{x_k\}$满足

$$\|x_{2k+1}-x^*\|_\infty=\|x_k-x^*\|_\infty, \tag{9.4.31}$$

$$\|x_{2k+2}-x^*\|_\infty=\|x_k-x^*\|_\infty^2. \tag{9.4.32}$$

所以双边既约海色阵方法不可能有一步 Q-超线性收敛. 事实上，由(9.4.31)可知连 Q-线性收敛性也没有.

从 $x_{k+1}=x_k+d_k$ 和(9.4.22)有

$$x_{k+1}=x_k-\begin{bmatrix}B_kZ_k^T\\ A_k^T\end{bmatrix}^{-1}\begin{bmatrix}Z_k^Tg_k\\ c_k\end{bmatrix}. \tag{9.4.33}$$

考虑 $B_k=B(x_k)$ 且 $B(x)$ 是连续可微的，于是我们可理解双边既约海色阵方法等价于不动点问题

$$x=\phi(x) \tag{9.4.34}$$

的简单迭代法，其中

$$\phi(x)=x-\begin{bmatrix}B(x)Z(x)^T\\ A(x)^T\end{bmatrix}^{-1}\begin{bmatrix}Z(x)^Tg(x)\\ c(x)\end{bmatrix}. \tag{9.4.35}$$

显然有

$$\begin{aligned}\phi(x)&=x-[Z(x)B(x)^{-1}\quad A(x)[A(x)^TA(x)]^{-1}]\\&\quad\times\begin{bmatrix}Z(x)^Tg(x)\\ c(x)\end{bmatrix}\\&=x-Z(x)B(x)^{-1}Z(x)^T[g(x)-A(x)\lambda^*]\\&\quad-A(x)[A(x)^TA(x)]^{-1}c(x).\end{aligned} \tag{9.4.36}$$

由于 $g(x^*)-A(x^*)\lambda^*=0$ 且 $c(x^*)=0$，从(9.4.36)式可知，

$$\nabla\phi(x^*)=I-Z(x^*)B(x^*)^{-1}Z(x^*)^TW(x^*,\lambda^*)$$
$$-A(x^*)[A(x^*)^TA(x^*)]^{-1}A(x^*)^T, \tag{9.4.37}$$

利用

$$A(x^*)=Q(x^*)\begin{bmatrix}R(x^*)\\0\end{bmatrix}$$
$$=[Y(x^*)\quad Z(x^*)]\begin{bmatrix}R(x^*)\\0\end{bmatrix} \tag{9.4.38}$$

以及记号 $W^*=W(x^*,\lambda^*)$ 可将(9.4.37)写成

$$\nabla\phi(x^*)=Q(x^*)$$
$$\times\begin{bmatrix}0 & 0\\ -B(x^*)^{-1}Z(x^*)^TW^*Y(x^*) & I-B(x^*)^{-1}Z(x^*)^TW^*Z(x^*)\end{bmatrix}$$
$$\times Q(x^*)^T. \tag{9.4.39}$$

令 $\rho(G)$ 为矩阵 G 的谱半径，且定义

$$\rho_1=\rho(R(x^*)^{-1}Z(x^*)^TW^*Y(x^*)), \tag{9.4.40}$$
$$\rho_2=\rho(I-B(x^*)^{-1}Z(x^*)^TW^*Z(x^*)). \tag{9.4.41}$$

则由(9.4.39)可知

$$\rho(\nabla\phi(x^*))\leqslant\rho_1+\rho_2 \tag{9.4.42}$$

以及

$$\rho([\nabla\phi(x^*)]^2)\leqslant\rho_2(\rho_1+\rho_2). \tag{9.4.43}$$

所以，只要 $\rho_2(\rho_1+\rho_2)<1$，则 x_k 局部两步线性收敛于 x^*. 特别地，$B(x^*)=Z(x^*)^TW^*Z(x^*)$，我们有 $\rho_2=0$，于是

$$\rho([\nabla\phi(x^*)]^2)=0. \tag{9.4.44}$$

故点列 x_k 两步超线性收敛于 x^*. 只要 $Z(x^*)^TW^*Y(x^*)\neq0$，则

$$\rho(\nabla\phi(x^*))\geqslant\rho_1\geqslant\frac{\rho(Z(x^*)^TW^*Y(x^*))}{\rho(B(x^*))}>0. \tag{9.4.45}$$

这也就说明为什么双边既约海色阵方法不可能一步超线性收敛。从(9.4.39)还可知道，如果 $Z(x^*)^TW^*Y(x^*)=0$，且

$$B(x^*)=Z(x^*)^TW^*Z(x^*),$$

则点列 x_k 一步 Q-超线性收敛于 x^*.

§9.5 信 赖 域 法

信赖域法的基本思想是要求试探步 d_k 在信赖域之内，即满足

$$\|d_k\| \leqslant \Delta_k, \tag{9.5.1}$$

其中 $\|\cdot\|$ 是 $\mathbf{R}^n$ 中某一范数. 约束优化应用信赖域法的一个主要困难是线性化约束条件(9.3.2)与(9.3.3)在信赖域 $\{d: \|d\| \leqslant \Delta_k\}$ 内无解. 对这一困难至少有两种处理方法：第一种方法是求解如下的子问题:

$$\min_{d \in \mathbf{R}^n} g_k^T d + \frac{1}{2} d^T B_k d, \tag{9.5.2}$$

$$\text{s. t.} \quad \theta c_i(x_k) + d^T a_i(x_k) = 0, \ i \in E, \tag{9.5.3}$$

$$\theta c_i(x_k) + d^T a_i(x_k) \geqslant 0, \ i \in I, \tag{9.5.4}$$

$$\|d\|_\infty \leqslant \Delta_k. \tag{9.5.5}$$

其中 $\theta \in (0, 1]$. 为了使 θ 尽可能靠近 1, 可在(9.5.2)的目标函数中加上一项 $\sigma(\theta-1)^2$. 关于这类方法的讨论可见 Byrd、Schnabel 和 Shultz(1987) 和 Calamai 和 Moré(1987). 另一种方法，仅为处理等式约束问题. 是考虑下面的子问题:

$$\min_{d \in \mathbf{R}^n} g_k^T d + \frac{1}{2} d^T B_k d, \tag{9.5.6}$$

$$\sum_{i=1}^{m} (c_i(x_k) + a_i(x_k)^T d)^2 \leqslant \xi_k, \tag{9.5.7}$$

$$\|d\|_2 \leqslant \Delta_k, \tag{9.5.8}$$

其中 $\xi_k \geqslant 0$ 可由某种方式求得[见 Celis、Dennis 和 Tapia(1985) 以及 Powell 和 Yuan(1991)].

在 Powell 和 Yuan(1991)提出的算法中，第 k 次迭代时的子问题为(9.5.6)～(9.5.8)，且 ξ_k 满足

$$\min_{\|d\|_2 \leqslant b_1 \Delta_k} \|c_k + A_k^T d\|_2^2 \leqslant \xi_k \leqslant \min_{\|d\|_2 \leqslant b_2 \Delta_k} \|c_k + A_k^T d\|_2^2, \tag{9.5.9}$$

其中 $b_1 \geqslant b_2$ 是(0, 1)区间上的两个给定常数. 线搜索用的罚函数是光滑精确罚函数

$$\phi_k(x)=f(x)-\lambda(x)^T c(x)+\sigma_k\|c(x)\|_2^2, \qquad (9.5.10)$$

其中 $\sigma_k>0$ 是当前的罚因子，$\lambda(x)$是最小二乘

$$\min_{\lambda\in\mathbf{R}^m}\|g(x)-A(x)\lambda\|_2 \qquad (9.5.11)$$

的最小范数解. 为了避免计算二阶导数，我们选取

$$\begin{aligned}D_k= & -(g_k-A_k\lambda_k)^T d_k-\frac{1}{2}d_k^T B_k\hat{d}_k \\ & -(\lambda(x_k+d_k)-\lambda(x_k))^T\left(c_k+\frac{1}{2}A_k^T d_k\right) \\ & -\sigma_k(\|c_k+A_k^T d_k\|_2^2-\|c_k\|_2^2)\end{aligned} \qquad (9.5.12)$$

作为价值函数 $\phi_k(x)$的预估下降量. 其中

$$\hat{d}_k=(I-A_kA_k^+)d_k \qquad (9.5.13)$$

是 d_k 投影到 A_k^T 的零空间上的向量. 然后计算 $\phi_k(x)$ 的真实下降量与预估下降量的比值

$$r_k=\frac{\phi_k(x_k)-\phi_k(x_k+d_k)}{D_k} \qquad (9.5.14)$$

下一个迭代点 x_{k+1} 的选取以及如何调节信赖域半径都和比值 r_k 有关. 下面的算法是由 Powell 和 Yuan(1991)给出的.

算法 9.5.1

步 1　给出 $x_1\in\mathbf{R}^n$, $B_1\in\mathbf{R}^{n\times n}$, $\Delta_1>0$, $0<b_2\leqslant b_1<1$;
　　$\sigma_1>0$, $\varepsilon\geqslant 0$, $k:=1$.

步 2　如果 $\|c_k\|_2+\|g_k-A_k\lambda_k\|\leqslant\varepsilon$, 则停;
　　解(9.5.6)～(9.5.8), 给出d_k;

步 3　由(9.5.12)计算 D_k;
　　如果

$$D_k\geqslant\frac{1}{2}\sigma_k[\|c_k\|_2^2-\|c_k+A_k^T d_k\|_2^2]; \qquad (9.5.15)$$

　　则转步 4
　　计算

$$\bar{\sigma}_k=2\sigma_k+\max\left\{0,\ \frac{-2D_k}{\|c_k\|_2^2-\|c_k+A_k^T d_k\|_2^2}\right\}, \qquad (9.5.16)$$

　　令　$D_k:=D_k+[\bar{\sigma}_k-\sigma_k][\|c_k\|_2^2-\|c_k+A_k^T d_k\|_2^2]$;

$\sigma_k := \bar{\sigma}_k$;

步 4 计算比值(9.5.14);

$$x_{k+1}=\begin{cases} x_k+d_k, & \text{如果 } r_k>0; \\ x_k, & \text{否则}. \end{cases} \tag{9.5.17}$$

$$\Delta_{k+1}=\begin{cases} \max[\Delta_k,\ 4\|d_k\|_2], & r_k>0.9; \\ \Delta_k, & 0.1\leqslant r_k\leqslant 0.9; \\ \min[\Delta_k/4,\ \|d_k\|_2/2], & r_k<0.1. \end{cases} \tag{9.5.18}$$

步 5 计算 B_{k+1}; $\sigma_{k+1} := \sigma_k$;

$k := k+1$; 转步 2.

对于等式约束问题(9.1.1)与(9.1.2), Powell 和 Yuan (1991)证明了如下收敛性定理:

定理 9.5.2 设 $\{x_k\}$、$\{d_k\}$、$\{B_k\}$一致有界, $A(x)$ 对所有 x 都是列满秩, 则必有

$$\lim_{k\to\infty}\inf[\|c_k\|_2+\|(I-A_kA_k^+)g_k\|_2]=0, \tag{9.5.19}$$

故知$\{x_k\}$必存在一个聚点为原问题(9.1.1)与(9.1.2)的稳定点.

定理 9.5.3 在定理 9.5.2 的假定下, 如果 $x_k\to x^*$, 在 x^* 处二阶充分条件满足, 且有

$$\lim_{k\to\infty}\max_{A_k^T d=0,\|d\|_2\leqslant 1}\frac{|d^T(B_k-W(x^*,\ \lambda^*))d|}{\|d_k\|_2}=0, \tag{9.5.20}$$

则有

$$\lim_{k\to\infty} r_k=1, \tag{9.5.21}$$

且 x_k 必 Q-超线性收敛于 x^*.

关于以上两个定理的证明可见 Powell 和 Yuan(1991).

求解约束优化的信赖域法的子问题主要为两类, 即(9.5.2)~(9.5.5)和(9.5.6)~(9.5.8). 对于前者, 子问题可化为一个二次规划, 故可用第六章讨论的方法求解. 下面讨论子问题(9.5.6)~(9.5.8).

考虑两个凸二次约束的求二次函数极小的问题:

$$\min_{d\in\mathbf{R}^n} g^Td+\frac{1}{2}d^TBd, \tag{9.5.22}$$

$$\text{s.t.}\quad \|A^Td+c\|_2\leqslant\xi, \tag{9.5.23}$$

$$\|d\|_2\leqslant\Delta. \tag{9.5.24}$$

其中 $g\in\mathbf{R}^n$, $B\in\mathbf{R}^{n\times n}$, $A\in\mathbf{R}^{n\times m}$, $c\in\mathbf{R}^m$, $\Delta>0$, $\xi\geqslant0$ 以及 B 是对称阵. 由于在

$$\xi=\xi_{\min}=\min_{\|d\|_2\leqslant\Delta}\|A^Td+c\|_2 \tag{9.5.25}$$

时,(9.5.23)与(9.5.24)在 $\|(A^T)^+c\|_2\geqslant\Delta$ 时只有一个可行点, 而在 $\|(A^T)^+c\|_2<\Delta$ 时, 可将(9.5.22)~(9.5.24)等价地化成

$$\min_{\bar d\in\mathbf{R}^n}(g-B(A^T)^+c)^T(I-(A^T)^+A^T)\bar d$$

$$+\frac{1}{2}\bar d^T(I-(A^T)^+A^T)B(I-(A^T)^+A^T)\bar d, \tag{9.5.26}$$

$$\text{s.t.}\quad \|\bar d\|_2\leqslant\sqrt{\Delta^2-\|(A^T)^+c\|_2^2}. \tag{9.5.27}$$

这变成在一个球形区域求二次函数的最小值. 关于它的讨论和算法可见 Gay(1981)以及 Moré 和 Sorensen(1983), 所以我们可以假定

$$\xi>\xi_{\min}. \tag{9.5.28}$$

Yuan(1990a)证明了如下结果:

定理 9.5.4 假定(9.5.28)成立, d^* 是问题(9.5.22)~(9.5.24)的解, 则必存在 Lagrange 乘子 $\lambda\geqslant0$, $\mu\geqslant0$, 使得

$$(B+\lambda I+\mu AA^T)d^*=-(g+\mu Ac), \tag{9.5.29}$$

且满足互补条件:

$$\lambda[\Delta-\|d^*\|_2]=0, \tag{9.5.30}$$

$$\mu[\xi-\|A^Td^*+c\|_2]=0, \tag{9.5.31}$$

而且如果乘子 λ、μ 是唯一确定的话, 矩阵

$$H(\lambda,\mu)=B+\lambda I+\mu AA^T \tag{9.5.32}$$

最多只有一个负特征值.

定理 9.5.5 在定理 9.5.4 的条件下, 记 Ω 为所有满足(9.5.29)~(9.5.31), 且 $\lambda\geqslant0$、$\mu\geqslant0$ 的乘子, 则必存在 $(\lambda,\mu)\in\Omega$ 使得矩阵(9.5.32)最多只有一个负特征值.

关于以上两个定理的证明可见 Yuan(1990a). 对于子问题

(9.5.22)～(9.5.24)至今尚无比较理想的求解方法．一种处理办法是只考虑二维子空间上的解，即考虑

$$d \in \text{span}\{g,\ (A^T)^+ c\}. \tag{9.5.33}$$

另一种方法，只对于 B 是正定的情形求解(9.5.22)～(9.5.24)的对偶问题：

$$\min\ \psi(\lambda,\ \mu), \tag{9.5.34}$$

$$\text{s.t.}\quad \lambda \geqslant 0,\ \mu \geqslant 0. \tag{9.5.35}$$

其中

$$\psi(\lambda,\ \mu) = \frac{1}{2}\lambda\Delta^2 + \frac{1}{2}\mu[\xi^2 - \|c\|_2^2] + \frac{1}{2}(g+\mu Ac)^T H(\lambda,\ \mu)^{-1}(g+\mu Ac). \tag{9.5.36}$$

且 $H(\lambda,\ \mu)$ 由(9.5.32)定义．引入记号

$$d(\lambda,\ \mu) = -H(\lambda,\ \mu)^{-1}(g+\mu Ac), \tag{9.5.37}$$

则不难求得

$$\nabla\psi(\lambda,\ \mu) = \frac{1}{2}\begin{bmatrix}\Delta^2 - \|d(\lambda,\ \mu)\|_2^2 \\ \xi^2 - \|c + A^T d(\lambda,\ \mu)\|_2^2\end{bmatrix}. \tag{9.5.38}$$

$$\nabla^2\psi(\lambda,\ \mu) = \begin{bmatrix} d(\lambda,\ \mu)^T H(\lambda,\ \mu)^{-1} d(\lambda,\ \mu) & d(\lambda,\ \mu)^T H(\lambda,\ \mu)^{-1} y(\lambda,\ \mu) \\ d(\lambda,\ \mu)^T H(\lambda,\ \mu)^{-1} y(\lambda,\ \mu) & y(\lambda,\ \mu)^T H(\lambda,\ \mu)^{-1} y(\lambda,\ \mu)\end{bmatrix}. \tag{9.5.39}$$

其中

$$y(\lambda,\ \mu) = A(c + A^T d(\lambda,\ \mu)). \tag{9.5.40}$$

于是可知，只要计算出 $H(\lambda,\ \mu)$ 的 Cholesky 分解，则很容易计算函数 $\psi(\lambda,\ \mu)$ 的一阶导数和二阶导数．所以我们可用牛顿法求解对偶问题(9.5.34)与(9.5.35)，详细的讨论可见 Yuan(1991)．一旦 $\lambda^* \geqslant 0$、$\mu^* \geqslant 0$ 为(9.5.34)与(9.5.35)的解已求出，我们令 $d^* = d(\lambda^*,\ \mu^*)$，即知 d^* 是原问题(9.5.22)～(9.5.24)的解．

第 10 章

非光滑优化

非光滑优化 (nonsmooth optimization) 是指函数 $f(x)$ 和 $c_i(x)$ $(i=1, \cdots, m)$ 不全是可微函数的优化问题 (1.1.1)～(1.1.3). 非光滑优化也被称为不可微优化 (nondifferentiable optimization).

非光滑优化是一个很大的问题，这方面的专著已不少(如见 Rockafellar (1970, 1981)、Clark (1983)等)，它决不是一章就能详细讨论的. 本章将主要讨论一类非常特殊形式的非光滑优化问题

$$\min_{x \in \mathbf{R}^n} h(f(x)) \tag{10.0.1}$$

以及求解它的信赖域方法. 在(10.0.1)中 $f(x)$ 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的连续可微函数. $h(\cdot)$ 是 $\mathbf{R}^m$ 上的凸函数. 这类特殊的非光滑优化问题被称为复合非光滑 (composite nonsmooth) 优化问题，或者复合不可微优化问题. 不可微优化也被简称为 NDO. 故问题 (10.0.1) 也可简称为复合 NDO 问题.

§10.1 方法概述

利用 L_1 精确罚函数，我们可把一个带约束的 NDO 问题在一定意义下等价地转化成一个无约束的 NDO，所以我们只讨论无约束的不可微优化问题

$$\min_{x \in \mathbf{R}^n} F(x), \tag{10.1.1}$$

这里我们对 $F(x)$ 不要求可微.

即使 $F(x)$ 几乎处处可微，利用解可微问题的方法来求解(10.1.1)(假定在每个迭代点上 $F(x)$ 均可微)将有两个很大的难点. 第一是算法的终止条件不易给出. 我们知道，当 x 充分靠近一连续可微函数 $f(x)$ 的极小点时 $\|\nabla f(x)\|$ 一定非常小. 所以，光滑的无约束优化方法的终止判别条件常常是 $\|\nabla f(x)\|\leqslant\varepsilon$. 但对于不可微函数并没有类似的结论. 例如 $n=1$, $F(x)=|x|$，则对任何不是解的 x 都有

$$|\nabla F(x)|=1.$$

另一个难点是由 Wolfe (1975)指出的"折线收敛于非解"现象，也就是说，当 $F(x)$ 是不可微函数时，精确搜索下的最速下降法可能收敛于一非稳定点(关于非光滑函数的稳定点可见 Rockafellaz (1981)以及本书的定义 10.2.1). 例如，设 $x=(u, v)^T$,

$$F(x)=\max\left\{\frac{1}{2}v^2+(u-1)^2,\ \frac{1}{2}v^2+(u+1)^2\right\}. \tag{10.1.2}$$

假定

$$x_k=\begin{bmatrix}2(1+|\varepsilon_k|)\\ \varepsilon_k\end{bmatrix}, \tag{10.1.3}$$

且 $\varepsilon_k\neq0$, 则不难求得精确搜索下的最速下降法将给出

$$x_{k+1}=\begin{bmatrix}2(1+|\varepsilon_k|/3)\\ -\varepsilon_k/3\end{bmatrix}. \tag{10.1.4}$$

于是，对于函数(10.1.2)以及给出的初始点 $(2+2|\delta|, \delta)^T$ $(\delta\neq0)$，精确线搜索下的最速下降法将收敛于 $(2, 0)^T$. 显然 $(2, 0)^T$ 不是(10.1.2)的稳定点.

非光滑优化的一类方法是次梯度(Subgradient)方法. 这类方法最早由 Shor (1962)给出. 对于凸函数 $F(x)$，我们称向量 g^* 为 $F(x)$ 在 x 点处的次梯度，当且仅当

$$F(z)\geqslant F(x)+(z-x)^Tg^*,\ \forall z\in\mathbf{R}^n. \tag{10.1.5}$$

$F(x)$ 在 x 处的所有次梯度组成的集合称为 $F(x)$ 在 x 处的次微分(subdifferential)，用 $\partial F(x)$ 表示之. 次梯度和次微分概念可

推广到一般不可微函数(如见 Clark, 1975). 次梯度法是梯度法的直接推广, 它利用

$$x_{k+1}=x_k-\alpha_k g_k/\|g_k\|_2 \tag{10.1.6}$$

产生点列 $\{x_k\}$, 其中 $g_k\in\partial F(x_k)$, $\alpha_k>0$ 是一步长因子. 从例子(10.1.2)可知 α_k 如果按精确搜索选取, 可能导致于收敛到非稳定点. 幸好的是: 如果 $F(x)$ 是凸函数且 x_k 不是解, 则对充分小的 $\alpha_k>0$, 必有

$$\|x_{k+1}-x^*\|_2<\|x_k-x^*\|_2. \tag{10.1.7}$$

其中 x^* 是(10.1.1)的任一极小点(见 Zowe, 1985). Shor (1962) 证明了对于凸函数 $F(x)$, 如果集合 $S=\{x: \operatorname{Arg}\min F(x)\}$ 非空, 则对任何 $\delta>0$, 存在 $\bar{r}>0$, 如果 $\alpha_k\equiv r\quad(r\in(0,\bar{r}))$, 则有

$$\liminf_{k\to\infty} F(x_k)\leqslant\min_{x\in\mathbf{R}^n} F(x)+\delta. \tag{10.1.8}$$

如果步长 $\alpha_k>0$ 满足

$$\lim_{k\to\infty}\alpha_k=0,\quad \sum_{k=1}^{\infty}\alpha_k=+\infty, \tag{10.1.9}$$

则可保证对凸函数 $F(x)$, 有

$$\lim_{k\to\infty} F(x_k)=\min_{x\in\mathbf{R}^n} F(x). \tag{10.1.10}$$

于是任何聚点都是极小点[见 Ermoliev(1966)和 Polyak(1967)], 由于(10.1.9)将使得算法收敛速度比 R-线性还要慢, Shor(1968) 建议取

$$\alpha_k=\alpha_1 q^k,\ 0<q<1. \tag{10.1.11}$$

从而有 R-线性收敛速度. 但选取方法(10.1.11)并不能保证收敛到解. 事实上, 只要 x_1 到 S 的距离大于 $\alpha_1/(1-q)$ 时, $\{x_k\}$ 则不可能靠近(10.1.1)的任何解.

如果知道目标函数的最小值 F^*, 则可选取

$$\alpha_k=\lambda\frac{F(x_k)-F^*}{\|g_k\|_2},\ 0<\lambda<2. \tag{10.1.12}$$

这一步长选取方法最初由 Eremin (1965)对于极大极小(minimax)问题

$$\min F(x)=\max_{1\leqslant i\leqslant m}\{f_i(x),\ 0\} \tag{10.1.13}$$

提出来的. Polyak (1969) 将这一方法推广到一般非光滑优化, 且证明了其 R-线性收敛性. 值得指出的是: 对于问题(10.1.13), 该方法等价于求解非线性不等式组

$$f_i(x)\leqslant 0,\ i=1,\ \cdots,\ m \tag{10.1.14}$$

的松弛(relaxation)方法. 松弛方法至少可溯源到 Agmon (1954)以及 Motzkin 和 Schoenberg(1954).

另一种加快次梯度法收敛速度的方法是 Shor (1970) 提出的空间扩充(space dilation)方法. 该方法的基本思想是: 在第 k 次迭代时, 搜索方向不仅和 g_k 有关, 而且也依赖于 g_{k-1}. 关于这方法的详细讨论可见 Goffin(1978), Polyak(1978), Skokov(1974), Shor(1977, 1983)以及 Zowe(1985).

解凸规划问题的割平面(cutting plane)法分别由 Kelley (1959)与 Cheney 和 Goldstein (1959)独立提出的. 割平面法的基本思想是: 每次迭代求函数在一凸多面体(polyhedron)的极小值. 每次迭代后引进一割平面, 从而缩小多面体. 设在 k 次迭代有多面体 S_k, 则割平面法求解

$$\min_{v\in\mathbb{R}^n} v \tag{10.1.15}$$

$$\text{s. t. } v\geqslant F(x_i)+g_i^T(x-x_i),\ i=1,\ \cdots,\ k, \tag{10.1.16}$$

$$x\in S_k. \tag{10.1.17}$$

当函数 $F(x)$可微时, 且假定算法收敛于解, 则对所有充分大的 k, $g_k=\nabla F(x_k)$非常小. 从而 (10.1.16) 将会是非常坏条件的(ill-conditioned). 割平面法的另一个缺点是: 当 k 充分大时, 问题(10.1.15)~(10.1.17)中约束太多. 正因为这些原因, 割平面法尽管是最早的解一般凸规划问题的方法, 它从未受到很大重视.

Wolfe (1975)提出共轭次梯度(conjugate subgradient)法. 其基本思想是将共轭梯度法推广到非光滑问题. 在第 k 次迭代, 定义一个集合 $I_k\subset\{1,\ 2,\ \cdots,\ k\}$, 求解

$$\min \Big\|\sum_{i\in I_k}\lambda_i g_i\Big\|_2^2, \tag{10.1.18}$$

$$\text{s. t. } \sum_{i\in I_k}\lambda_i=1,\ \lambda_i\geqslant 0, \tag{10.1.19}$$

设 $\lambda_i^{(k)}$ $(i\in I_k)$是(10.1.18)与(10.1.19)的解. 然后取

$$d_k=-\sum_{i\in I_k}\lambda_i^{(k)}g_i \tag{10.1.20}$$

为第 k 次迭代的线搜索方向. 设 $I_{k+1}=I_k\cup\{k+1\}$, 则当 $F(x)$是凸的二次函数时, 精确线搜索下的共轭次梯度法实质上就是共轭梯度法.

共轭次梯度法的推广是捆集(bundle)法. 捆集法在第 k 次迭代时有加权因子 $t_i^{(k)}$ $(i=1, \cdots, k)$. 且求解

$$\min \left\|\sum_{i=1}^{k}\lambda_i g_i\right\|_2^2, \tag{10.1.21}$$

$$\text{s. t. } \sum_{i=1}^{k}\lambda_i=1,\ \lambda_i\geqslant 0, \tag{10.1.22}$$

$$\sum_{i=1}^{k}\lambda_i t_i^{(k)}\leqslant s. \tag{10.1.23}$$

其中 $s>0$ 是一预先给定的常数. 设 $\lambda_i^{(k)}$ $(i=1, \cdots, k)$是(10.1.21)~(10.1.23)的解. 捆集法取

$$d_k=-\sum_{i=1}^{k}\lambda_i^{(k)}g_i \tag{10.1.24}$$

为第 k 次迭代的搜索方向. 不难看出, 如果 $t_i^{(k)}=0$ $(i\in I_k)$且 $t_i^{(k)}=+\infty$ $(i\notin I_k)$, 则(10.1.18)~(10.1.20)和(10.1.21)~(10.1.24)完全等价. 关于捆集法的详细讨论以及它和次梯度法和割平面法的关系可见 Lemaréchal (1978a、b, 1980), Lemarechal, Strodiot 和 Bihain (1981)和 Zowe (1985).

Fletcher (1981) 给出了一个求解复合 NDO 问题的信赖域法. 关于复合 NDO 的信赖域法的研究很多, 如 Burke (1985), Fletcher (1982a, b), Womersley (1985)和 Yuan (1983a, b, 1984b, 1985a、b)等. 我们要在本章后面的几节详细讨论.

其他还有不少关于非光滑优化的工作, 感兴趣的读者可参阅 Balinski 和 Wolfe(1975), Demyanov 和 Dixon(1986), Dixon, Spedicato 和 Szego(1980), Lemarechal 和 Mifflin(1978), Nurminski(1982), 以及 Sorensen 和 Wets(1982)等文集.

§10.2　复合NDO的基本性质

从现在起，我们仅考虑复合 NDO 问题

$$\min_{x\in \mathbf{R}^n} h(f(x)), \tag{10.2.1}$$

其中 $f(x)=(f_1(x), \cdots, f_m(x))^T$ 是连续可微函数，$h(\cdot)$是 $\mathbf{R}^m$ 中的凸函数.

复合NDO问题在离散逼近以及数据拟合等方面常常可见到，如见 Charalambous 和 Bandler (1976)、Charalambous 和 Conn (1975, 1978), Hald 和 Madsen (1979, 1981), Hornung (1982 a、b, 1983), Madsen (1975), Murray 和 Overton (1980, 1981) Osborne 和 Watson(1969), Powell (1984c) 以及 Watson(1978, 1979)等. 下面的简单例子是将一个逼近问题化成一个非光滑优化问题.

考虑线性方程组

$$Ax=b, \tag{10.2.2}$$

其中 $A\in \mathbf{R}^{m\times n}$, $b\in \mathbf{R}^m$. 当 $m>n$ 时，方程(10.2.2)一般无解. 所以我们可取 x, 使得 Ax 和 b 之间的误差尽可能小，也就是说我们需要求解

$$\min_{x\in \mathbf{R}^n} \|Ax-b\|, \tag{10.2.3}$$

其中 $\|\cdot\|$ 是 $\mathbf{R}^m$ 中的某一范数. 不难看出，(10.2.3)具有形式(10.2.1). 在(10.2.3)中令 $\|\cdot\|$ 为欧氏范数，则问题是经典的最小二乘(least square)问题.

复合 NDO 问题越来越吸引人们注意的另一个原因是：一般的约束规划问题可通过 L_1 精确罚函数改写成一个复合 NDO 问题.

下面我们考虑复合NDO问题的最优性条件. 为了叙述简单，我们引入如下记号：

$$\chi(x, d)=h(f(x))-h(f(x)+(\nabla f(x)^T)^T d), \tag{10.2.4}$$

$$\psi_t(x)=\max_{\|d\|\leqslant t}\{\chi(x,d)\},\tag{10.2.5}$$

$$DF(x,d)=\sup_{\lambda\in\partial h(f(x))}d^T\nabla f(x)^T\lambda,\tag{10.2.6}$$

其中 $\partial h(f(x))$ 是指函数 $h(\cdot)$ 在 $f(x)$ 点处的次微分.

定义 10.2.1 如果

$$DF(x^*,d)\geqslant 0,\ \forall d\in\mathbf{R}^n,\tag{10.2.7}$$

则称 x^* 是函数 $h(f(x))$ 的稳定点.

从 Rockafellar (1970, 1981) 即知下面结果成立:

引理 10.2.2 1) $DF(x,d)$ 对一切 x 和 d 都存在;

2) $\chi(x,d)$ 看成是 x 的函数时是一个凸函数; $\chi(x,d)$ 看成是 d 的函数时是一个凹函数, 且在 $d^*=0$ 处沿方向 d 的方向导数为 $DF(x,d)$;

3) $\psi_t(x)\geqslant 0,\ \forall t\geqslant 0$; $\psi_t(x)=0$ 当且仅当 x 是 $h(f(x))$ 的稳定点;

4) $\psi_t(x)$ 是 t 的凹函数;

5) $\psi_t(x)$ 对任何给定的 $t\geqslant 0$ 都是 x 的连续函数.

利用上述结果, 我们可证 $\{x_k\}$ 存在一个为稳定点的聚点 x^* 等价于

$$\liminf_{k\to\infty}\psi_1(x_k)=0.\tag{10.2.8}$$

下面的引理给出一阶的必要条件:

引理 10.2.3 如果 x^* 是复合 NDO 问题(10.2.1)的局部极小点, 则 x^* 必是 $h(f(x))$ 的稳定点.

证明 对任给的 $d\in\mathbf{R}^n$, 由于 x^* 是(10.2.1)的极小点,

$$h(f(x^*+\alpha d))\geqslant h(f(x^*))\tag{10.2.9}$$

对所有充分小的 α 都成立. 因为 $h(\cdot)$ 在整个 $\mathbf{R}^m$ 上有定义且是凸的, 故它必连续(见 Rockafellar, 1970). 从而由(10.2.9)可知:

$$h(f(x^*)+\alpha(\nabla f(x^*)^T)^Td)+o(\alpha)\geqslant h(f(x^*)).\tag{10.2.10}$$

于是有

$$\liminf_{\alpha\to 0_+}\frac{h(f(x^*)+\alpha(\nabla f(x^*)^T)^Td)-h(f(x^*))}{\alpha}\geqslant 0. \tag{10.2.11}$$

由次梯度的定义和(10.2.11)可证

$$\max_{\lambda\in\partial h^*}\lambda^T(\nabla f(x^*)^T)^Td\geqslant 0, \tag{10.2.12}$$

其中 $\partial h^*=\partial h(f(x^*))$. 由于 d 的任意性, 从(10.2.12)即知(10.2.7)成立. 所以引理为真. ∎

复合 NDO 的最优性条件可写成不同的形式. 例如可把一阶最优性条件(10.2.7)写成下面的等价形式:

引理 10.2.4 如果 x^* 是复合 NDO 问题(10.2.1)的局部极小点, 则必存在 $\lambda^*\in\partial h^*$, 使得

$$\nabla f(x^*)^T\lambda^*=0. \tag{10.2.13}$$

证明 显然, 我们只需证明(10.2.13)和(10.2.7)等价就足够了. 不难看出, 由(10.2.13)可推出(10.2.7).

现假定(10.2.7)成立. 如果(10.2.13)不成立, 则集合

$$S=\{\nabla f(x^*)^T\lambda\mid\lambda\in\partial h^*\} \tag{10.2.14}$$

不包含原点. 由于 ∂h^* 是凸集, 所以 S 也是凸集. 于是必存在向量 $\bar d$, 使得 $\bar d$ 是 S 的法锥(normal cone)的内点. 从而有

$$\bar d^T\nabla f(x^*)^T\lambda<0,\ \forall\lambda\in\partial h^*. \tag{10.2.15}$$

由于 ∂h^* 是闭集, (10.2.15)显然与 $DF(x^*,\bar d)\geqslant 0$ 相矛盾. 此矛盾说明了(10.2.13)和(10.2.7)等价. ∎

对于光滑优化问题不可能有一阶充分条件, 但对于复合 NDO 问题有如下充分性结果:

引理 10.2.5 设对一切非零向量 $d\in\mathbf{R}^n$ 都有

$$DF(x^*,d)>0, \tag{10.2.16}$$

则 x^* 必是函数 $h(f(x))$ 的局部严格极小点.

证明 由(10.2.16)可知存在 $\delta>0$, 使得

$$DF(x^*,d)\geqslant\delta,\ \forall\|d\|_2=1. \tag{10.2.17}$$

如果引理不真, 则存在 $x_k\to x^*$, 且 $h(f(x_k))\leqslant h(f(x^*))$. 不妨设 $x_k=x^*+\alpha_kd_k$, $\|d_k\|_2=1$, $\alpha_k>0$ 且 $\alpha_k\to 0_+$. 于是

$$\begin{aligned}&h(f(x_k))-h(f(x^*))\\&=h(f(x^*)+(\nabla f(x^*)^T)^T(x_k-x^*))-h(f(x^*))+o(\alpha_k)\\&\geqslant\alpha_k DF(x^*,d_k)+o(\alpha_k)\\&\geqslant\alpha_k\delta+o(\alpha_k).\end{aligned}\tag{10.2.18}$$

这与 $h(f(x_k))\leqslant h(f(x^*))$ 相矛盾. 故知引理成立. ∎

对于非光滑问题的高阶最优性条件将依赖 $h(\cdot)$ 的广义高阶导数. 对于比较简单的 $h(\cdot)$, 如分片线性函数(polyhedral):

$$h(f)=\max_{1\leqslant i\leqslant l}(h_i^T f+b_i),\tag{10.2.19}$$

其中 $h_i\in\mathbf{R}^m, b_i\in\mathbf{R}^1$, 则(10.2.1)的二阶最优性条件是十分容易导出的. 如见 Fletcher 和 Watson (1980). 对于一般凸的复合 NDO 问题二阶最优性条件可见 Burke (1987).

§10.3 信赖域法

信赖域法是一类迭代算法, 给出初值 $x_1\in\mathbf{R}^n$, 信赖域法按下列方式产生点列 $\{x_k\}$. 在第 k 次迭代, 有当前迭代点 x_k, 信赖域半径 $\Delta_k>0$ 以及一 $n\times n$ 对称阵 B_k. 信赖域法求解子问题

$$\min_{d\in\mathbf{R}^n}\phi_k(d)=h(f(x_k)+(\nabla f(x_k)^T)^Td)+\frac{1}{2}d^TB_kd,\tag{10.3.1}$$

$$\text{s. t. }\|d\|\leqslant\Delta_k.\tag{10.3.2}$$

这里 $\|\cdot\|$ 是 $\mathbf{R}^n$ 中的某一给定范数. 由于欧氏空间的任何两个范数都相互等价, 以后除特别声明外, 我们均假定 $\|\cdot\|$ 就是 $\|\cdot\|_2$. (10.3.1)与(10.3.2)的解用 d_k 表示, 且称为第 k 次迭代的试探步(trial step). 下一次迭代的迭代点由下列给出:

$$x_{k+1}=\begin{cases}x_k+d_k, & \text{如果 } F(x_k)>F(x_k+d_k);\\ x_k, & \text{否则}.\end{cases}\tag{10.3.3}$$

这里我们用 $F(x)$ 表示 $h(f(x))$.

值得指出的是: 我们选取 x_{k+1} 的方式和Powell (1984c)的方式不一样. 我们的接受 $x_{k+1}=x_k+d_k$ 的条件要弱, 从而更容易取

$x_{k+1}=x_k+d_k$. 公式(10.3.3)的一个很好的性质是：只要新点使目标函数值下降就被接受. 有些算法要求目标函数值能"足够下降"才接受新点. 所谓"足够下降"，一般是指 d_k 满足类似于

$$r_k=\frac{F(x_k)-F(x_k+d_k)}{\phi_k(0)-\phi_k(d_k)} \tag{10.3.4}$$

不小于某一事先给定的正常数(见 Powell, 1984c). 我们偏向于(10.3.3)的原因是不接受一个较"好"的点(目标函数值较小)看来毫无道理.

(10.3.4)定义的 r_k 是目标函数的实际下降量(actual reduction)和预估下降量(predicted reduction)之比. 下一次迭代所用的信赖域半径 Δ_{k+1} 满足：

$$\|d_k\|\leqslant\Delta_{k+1}\leqslant\min[c_1\Delta_k,\ \bar{\Delta}],\ \text{如果}\ r_k\geqslant c_2, \tag{10.3.5}$$

$$c_3\|d_k\|\leqslant\Delta_{k+1}\leqslant c_4\Delta_k,\ \text{如果}\ r_k<c_2, \tag{10.3.6}$$

其中 c_i $(i=1, 2, 3, 4)$是正常数且满足 $c_1>1>c_4>c_3$, $c_2<1$; $\bar{\Delta}>0$ 是一预先给定的常数，它是所有信赖域半径的上界.

类似引理 10.2.4，我们可证必存在

$$\lambda_k\in\partial h(f(x_k)+(\nabla f(x_k)^T)^T d_k),\ \mu_k\in\partial\|d_k\|,\ \bar{\mu}_k\geqslant 0,$$

使得

$$\nabla f(x_k)^T\lambda_k+B_k d_k+\bar{\mu}_k\mu_k=0, \tag{10.3.7}$$

$$\bar{\mu}_k[\Delta_k-\|d_k\|]=0. \tag{10.3.8}$$

选定特定的 c_i 以及利用 $f(x)$ 的二阶导数，Fletcher (1981, 1982a)给出如下求解问题(10.2.1)的信赖域方法：

算法 10.3.1　步 1　给出 x_1, λ_0, $\Delta_1>0$, $\varepsilon\geqslant 0$; $k:=1$;

步 2　计算

$$B_k=\sum_{i=1}^{m}(\lambda_{k-1})_i\nabla^2 f_i(x_k). \tag{10.3.9}$$

求解(10.3.1)与(10.3.2)，给出 d_k.

如果 $\|d_k\|\leqslant\varepsilon$，则停；

步 3　计算 r_k;

如果 $r_k<0.25$，则令 $\Delta_{k+1}:=\|d_k\|/4$;

如果 $r_k>0.75$，且 $\|d_k\|=\Delta_k$，则 $\Delta_{k+1}=2\Delta_k$;

如果 Δ_{k+1} 还未定义，则令 $\Delta_{k+1}=\Delta_k$.

步 4　如果 $r_k>0$，则转步 5；

$x_{k+1}=x_k$，$\lambda_k:=\lambda_{k-1}$；转步 6；

步 5　$x_{k+1}=x_k+d_k$；λ_k 由(10.3.7)定义；

步 6　$k:=k+1$，转步 2.

我们现在分析算法的收敛性．假定算法 10.3.1 产生的点列 $\{x_k\}$有界，这一假定在集合 $\{x|h(f(x))\leqslant h(f(x_1))\}$ 有界时显然成立．于是存在有界凸闭集 D，使得

$$x_k\in D,\ x_k+d_k\in D,\ \forall k=1,\ 2,\ \cdots. \tag{10.3.10}$$

由于 $h(\cdot)$是凸函数且在整个空间 $\mathbf{R}^m$ 有定义，故必存在常数 $L>0$，使得

$$|h(f_1)-h(f_2)|\leqslant L\|f_1-f_2\| \tag{10.3.11}$$

对所有的 f_1、$f_2\in f(D)=\{v=f(x)|x\in D\}$都成立（见 Rockafellar, 1970，第 237 页）．另外，由于 $f(x)$连续可微和 D 的有界性，存在常数 $M>0$，使得

$$\|\nabla f(x)^T\|\leqslant M \tag{10.3.12}$$

对一切 $x\in D$ 都成立.

Fletcher(1981)利用下面的引理 10.3.2 证明了定理 10.3.3.

引理 10.3.2　设 S 是所有满足 $y_k\to x^*$，$\varepsilon_k\to 0$ 的序列组成的集合，则对所有的 $d\in\mathbf{R}^n$ 都有

$$\limsup_{S}\frac{F(y_k+\varepsilon_k d)-F(x_k)}{\varepsilon_k}=\max_{\lambda\in\partial h^*}d^T\nabla f(x^*)^T\lambda. \tag{10.3.13}$$

定理 10.3.3　设 $f(x)$ 是两次连续可微的，如果算法 10.3.1 产生的点列 $\{x_k,\ k=1,\ \cdots\}$ 有界，则 $\{x_k\}$ 必存在一个聚点 x^* 是稳定点.

关于引理 10.3.2 和定理 10.3.3 的证明可见 Fletcher(1981)，Fletcher (1981)还指出：如果 B_k 不是由(10.3.9)给出，但只要 $\|B_k\|$ 一致有界，则上面的收敛性结果仍成立．另外，值得强调的是：尽管 Fletcher 的证明是基于特定的 c_i $(i=1,\ 2,\ 3,\ 4)$，但从

证明的过程可看出这些常数的大小并不影响证明的关键，所以 Fletcher 的收敛结果对整类方法均成立.

我们现在把 $\|B_k\|$ 一致有界的条件放宽到

$$\|B_k\| \leqslant c_5 + c_6 \sum_{i=1}^{k} \Delta_i, \tag{10.3.14}$$

其中 c_5、c_6 是两个正常数.

引理 10.3.4 设 d_k 是(10.3.1)与(10.3.2)的解. 则必有

$$F(x_k) - \phi_k(d_k) \geqslant \frac{1}{2} \psi_{\Delta_k}(x_k) \min\left\{1, \frac{\psi_{\Delta_k}(x_k)}{\|B_k\| \Delta_k^2}\right\}. \tag{10.3.15}$$

其中 $\psi_t(x)$ 由(10.2.4)与(10.2.5)所定义.

证明 由 d_k 的定义

$$F(x_k) - \phi_k(d_k) \geqslant F(x_k) - \phi_k(d), \quad \forall \|d\| \leqslant \Delta_k. \tag{10.3.16}$$

由 $\psi_t(x)$ 的定义，存在 $\|\bar{d}_k\| \leqslant \Delta_k$，使得

$$\psi_{\Delta_k}(x_k) = F(x_k) - h(f(x_k) + (\nabla f(x_k)^T)^T \bar{d}_k). \tag{10.3.17}$$

于是，由 $h(\cdot)$ 的凸性和(10.3.16)可知

$$\begin{aligned} F(x_k) - \phi_k(d_k) &\geqslant F(x_k) - \phi_k(\alpha \bar{d}_k) \\ &= \chi(x_k, \alpha \bar{d}_k) - \frac{1}{2} \alpha^2 \bar{d}_k^T B_k \bar{d}_k \\ &\geqslant \alpha \chi(x_k, \bar{d}_k) - \frac{1}{2} \alpha^2 \|B_k\| \cdot \|\bar{d}_k\|^2 \\ &\geqslant \alpha \psi_{\Delta_k}(x_k) - \frac{1}{2} \alpha^2 \|B_k\| \Delta_k^2 \end{aligned} \tag{10.3.18}$$

对一切 $\alpha \in [0, 1]$ 均成立. 从而有

$$\begin{aligned} F(x_k) - \phi_k(d_k) &\geqslant \max_{0 \leqslant \alpha \leqslant 1} \left\{\alpha \psi_{\Delta_k}(x_k) - \frac{1}{2} \alpha^2 \|B_k\| \Delta_k^2\right\} \\ &\geqslant \frac{1}{2} \min\left\{\psi_{\Delta_k}(x_k), \frac{[\psi_{\Delta_k}(x_k)]^2}{\|B_k\| \Delta_k^2}\right\}. \end{aligned} \tag{10.3.19}$$

即知引理成立. ∎

利用上面的引理可将定理 10.3.3 推广到以下定理:

定理 10.3.5 设 $f(x)$ 是两次连续可微的，如果算法 10.3.1 中 B_k 不从(10.3.9)给出但满足(10.3.14)，假定算法所产生的点列 $\{x_k, k=1, 2, \cdots\}$ 有界，则 $\{x_k\}$ 必存在一个聚点为稳定点.

证明 假定定理不真，则必存在 $\delta>0$，使得

$$\psi_1(x_k)\geqslant\delta \tag{10.3.20}$$

对一切 k 均成立. 从引理 10.2.2 的 4)、引理 10.3.4、不等式(10.3.20)和 Δ_k 的有界性，我们可证

$$F(x_k)-\phi_k(d_k)\geqslant c_7\min\left\{\Delta_k,\ \frac{1}{\|B_k\|}\right\}. \tag{10.3.21}$$

其中 c_7 是一正常数. 定义集合

$$S=\{k\mid r_k\geqslant c_2\}, \tag{10.3.22}$$

于是有

$$\begin{aligned}F(x_1)-\min_{x\in D}F(x)&\geqslant\sum_{k=1}^{\infty}[F(x_k)-F(x_{k+1})]\\&\geqslant\sum_{k\in S}[F(x_k)-F(x_{k+1})]\\&\geqslant c_2\sum_{k\in S}[F(x_k)-\phi_k(d_k)].\end{aligned} \tag{10.3.23}$$

由(10.3.21)、(10.3.23)和(10.3.14)可证

$$\sum_{k\in S}\frac{\Delta_k}{c_5+c_6\sum_{i=1}^{k}\Delta_i}<+\infty. \tag{10.3.24}$$

从(10.3.6)可知

$$\Delta_{k+1}\leqslant c_4\Delta_k,\quad 如果\ k\notin S. \tag{10.3.25}$$

从而不难证明(见 Powell, 1975)

$$\sum_{i=1}^{k}\Delta_i\leqslant\left(1+\frac{c_1}{1-c_4}\right)\left[\Delta_1+\sum_{\substack{i=1\\ i\in S}}^{k}\Delta_i\right]. \tag{10.3.26}$$

利用(10.3.24)和(10.3.26)可证 $\sum_{k\in S}\Delta_k$ 收敛，于是从(10.3.26)可知 $\sum_{k=1}^{\infty}\Delta_k$ 收敛. 故知 $\|B_k\|$ 一致有界. 由定理 10.3.3 知(10.3.20)不可能对一切 k 都成立. 此矛盾说明定理为真. ∎

Yuan (1985a)将上面收敛性结果的条件进一步放宽了，即把条件(10.3.14)换成了

$$\|B_k\|\leqslant c_8+c_9k. \tag{10.3.27}$$

其中 c_8 和 c_9 是两个正常数. Yuan(1985a)证明了只要

$$\sum_{k=1}^{\infty}\frac{1}{M_k}=+\infty, \tag{10.3.28}$$

则必有收敛性结果

$$\liminf_{k\to\infty}\psi_1(x_k)=0, \tag{10.3.29}$$

其中

$$M_k=\max_{1\leqslant i\leqslant k}\|B_i\|+1, \tag{10.3.30}$$

这实质上是将Powell (1984d)的结果从光滑优化推广到了非光滑优化. 条件(10.3.28)不能再进一步放宽了(见 Powell, 1984d).

看来，$\|B_k\|$ 的有界性和算法的收敛性之间的关系是十分有趣的. 虽然 $\|B_k\|$ 的有界性是算法收敛的一个充分条件，但 $\|B_k\|$ 的有界对收敛甚至超线性收敛也不必要. 当然，由 Powell (1984c)的结果可知，如果点列 x_k 超线性收敛，则

$$|d_k^TB_kd_k|/\|d_k\|^2 \tag{10.3.31}$$

必有界.

关于 B_k 的修正，我们可令

$$s_k=d_k, \tag{10.3.32}$$

$$y_k=\sum_{i=1}^{m}(\lambda_k)_i[\nabla f_i(x_k+d_k)-\nabla f_i(x_k)], \tag{10.3.33}$$

然后用第四章介绍的拟牛顿公式进行修正，其中 λ_k 是满足(10.3.7)与(10.3.8)的乘子.

如果 $h(f)$是分片线性函数(10.2.19)，且(10.3.2)中的范数为 $\|\cdot\|_\infty$ 或 $\|\cdot\|_1$，则问题(10.3.1)与(10.3.2)可化为一个二次规划问题.

对于特殊情形 $h(f)=f$, $m=1$，讨论的方法退化到光滑无约束优化的信赖域法. 在这种特殊情形，在(10.3.2)中可取 $\|\cdot\|$ 为 $\|\cdot\|_2$，解(10.3.1)与(10.3.2)的方法可见 Gay (1981) 和 Moré 和

Sorensen (1983).

§10.4 线性收敛的例子

本节给出例子说明由于目标函数的不可微性，上节给出的信赖域法的收敛速度可能仅为线性的.

由于我们仅考虑收敛速度，在本节中假定由算法产生的点列 x_k 收敛于问题(10.2.1)的一个解.

Powell (1975) 对于光滑情形 ($h(f)=f$, $m=1$) 证明了如果 $\nabla^2 f(x^*)$正定，且

$$\|(B_k-\nabla^2 f(x^*))d_k\|=o(\|d_k\|), \tag{10.4.1}$$

则信赖域方法必超线性收敛. 对于非光滑优化，这个结果不成立，下面给出的极大极小问题表明信赖域法可能仅是线性收敛.

为了使我们的例子对更多的算法均成立，我们假定算法定义

$$x_{k+1}=\begin{cases} x_k+d_k, & \text{如果 } r_k>c_0; \\ x_k, & \text{否则}. \end{cases} \tag{10.4.2}$$

其中 $c_0\in[0, c_2)$是一常数. 不难看出 (10.3.3) 就是 (10.4.2) 当 $c_0=0$ 时的特殊情形. 这里 r_k 的定义仍由(10.3.4)给出. 通常，信赖域法满足以下条件：

条件 10.4.1

1) $c_3\|d_k\|\leqslant\Delta_{k+1}\leqslant c_1\Delta_k$;

2) 如果 $r_k<c_2$, 则 $\Delta_{k+1}\leqslant c_4\Delta_k$;

3) 如果 $r_k\geqslant c_2'$, 则 $\Delta_{k+1}\geqslant\|d_k\|$.

其中 c_i ($i=1, 2, 3, 4$)和上一节给出的相同，即满足 $c_1>1>c_4>c_3$, $c_2<1$ 的正常数; c_2' 是满足 $c_2\leqslant c_2'<1$ 的常数. 具体实用的信赖域法常修正信赖域半径 Δ_{k+1}, 使其满足以下条件：

条件 10.4.2

1) 如果 $r_k\geqslant c_2'$ 且 $\|d_k\|=\Delta_k$, 则 $\Delta_{k+1}=c_1\Delta_k$;

2) 如果 $r_k\in[c_2, c_2')$, 则 $\Delta_{k+1}=\Delta_k$;

3) 如果 $r_k<c_2$, 则 $\Delta_{k+1}=c_4\Delta_k$ 或 $\Delta_{k+1}=c_4\|d_k\|$.

满足上述条件的信赖域法是很多的，如见 Fletcher (1980), Moré (1983), Powell (1984d), Yuan (1983, a, b) 等. 所有这些信赖域法均要求条件 10.4.1 中的 2). 我们发现这一性质就足够导致信赖域法仅有线性收敛速度.

我们构造例子的基本思想是设法使

$$x_{k+1}=x_k+d_k, \tag{10.4.3}$$

$$r_k\in(c_0,\ c_2), \tag{10.4.4}$$

$$\|x_{k+1}-x^*\|=\alpha\|x_k-x^*\| \tag{10.4.5}$$

对一切 k 都成立，α 是一个小于 1 的正常数.

我们考虑极大极小问题

$$\min_x h(f)=\min_x\,[\max\{f_1,\ f_2\}], \tag{10.4.6}$$

其中 $x=(u,\ v)^T$

$$f_1=1+v-u^2, \tag{10.4.7}$$

$$f_2=1-v+(1+\varepsilon)u^2, \tag{10.4.8}$$

其中 $\varepsilon>0$ 是一个参数且满足 $0<\varepsilon\leqslant 1$. 不难看到 (10.4.6) 的解是 $x^*=(0,0)^T$, 把问题(10.4.6)写成一个约束优化问题(Overton, 1982), 我们可发现相应的 Lagrange 乘子

$$\lambda^*=\left(\frac{1}{2},\ \frac{1}{2}\right)^T.$$

而且在解处严格互补 (strict complementarity) 条件和二阶充分条件都满足，且有

$$G_\infty^*=\sum_{i=1}^{2}\lambda_i^*\nabla^2 f_i(u^*,\ v^*)=\begin{bmatrix}\varepsilon & 0\\ 0 & 0\end{bmatrix}. \tag{10.4.9}$$

于是我们在子问题(10.3.1)与 (10.3.2) 中令 $B_k=G_\infty^*$. 我们的动机是基于下面的结果：如果 $x_k\to x^*$, $x_{k+1}=x_k+d_k$, $d=d_k$ 是 (10.3.1)的局部极小点，以及 $\|B_k\|$ 一致有界，则 x_k 超线性收敛于 x^* 等价于

$$\lim_{k\to\infty}\frac{\|P_\infty^*(B_k-G_\infty^*)d_k\|}{\|d_k\|}=0, \tag{10.4.10}$$

其中 P_∞^* 是从 $\mathbb{R}^n$ 到子空间 $\{d:\ d^T\nabla f_i(x^*)=0,\ i=1,\ 2\}$ 的投影

算子(见 Powell (1984c)以及 Powell 和 Yuan(1984)).

我们构造的点列 $x_k=(u_k, v_k)^T$ 都在一条曲线 $v=\theta u^2$ 上, 其中 $\theta<1+\varepsilon/2$ 是一常数. 我们使 $u_{k+1}=\alpha u_k$, 取范数 $\|\cdot\|$ 为 $\|\cdot\|_\infty$. 从而有

$$\|x_{k+1}-x^*\|_\infty=|u_{k+1}|=\alpha|u_k|=\alpha\|x_k-x^*\|_\infty. \tag{10.4.11}$$

故得到仅线性收敛的例子. 从而有以下结论:

结论 10.4.3 对于任何满足式(10.4.2)和条件 10.4.2 中 3) 的信赖域法对任意给定的 $0\leqslant c_0<c_2<1$, 都存在 $c_4\in(0, 1)$ 使得算法仅是线性收敛的.

关于结论 10.4.3 的详细证明可见 Yuan (1984b)或者 Yuan (1985d). 利用同一个问题(10.4.6), Yuan (1984b) 还证明了 Fletcher 的信赖域法(即算法 10.3.1) 可能是线性收敛且有两步估计式:

$$\|x_{k+2}-x^*\|_\infty=\frac{1}{2}\|x_k-x^*\|_\infty, \tag{10.4.12}$$

而且不难看出单靠调节 c_4 的值, 是不能克服仅仅线性收敛这一现象的. 所以我们讨论的这一类信赖域法(即满足式(10.4.2) 和条件 10.4.1) 均可能出现线性收敛.

从 Yuan (1984b) 给出的例子还可看到, 信赖域界在收敛速度方面起了很大作用. 事实上, 对于许多信赖域方法, 信赖域约束不积极(即 $\|d_k\|<\Delta_k$)是算法超线性收敛的一个必要条件, 而且在一定条件下, 信赖域约束不积极也是超线性收敛的一个充分条件(如见 Fletcher (1980), Powell 和 Yuan (1984)). 在我们的极大极小例子(10.4.6)中, 除了每次迭代信赖域约束均积极外, Powell 和 Yuan (1984)给出的其他的超线性收敛条件均满足. 所以, 信赖域约束的积极性破坏了算法的超线性收敛性.

不难发现, 线性收敛的一个原因是目标函数的导数有间断点. 在光滑的情形($h(f)=f$, $m=1$), 我们有

$$f(x_k+d_k)=f(x_k)+d_k^T\nabla f(x_k)+\frac{1}{2}d_k^TB_kd_k+o(\|d_k\|^2), \tag{10.4.13}$$

其中假定 B_k 满足(10.4.1). 从而由(10.4.13)即可知 $r_k\to 1$. 进而可证明超线性收敛(Powell, 1975). 但是,对于非光滑问题,关系式

$$h(f(x_k+d_k))=h(f(x_k)+(\nabla f(x_k)^T)^T d_k)+\frac{1}{2}d_k^T B_k d_k + o(\|d_k\|^2) \tag{10.4.14}$$

一般不成立,即使 $B_k=G_\infty^*$ (10.4.14)也可能不真. 于是当 $\{x_k\}$ 沿着 $[h(f(x_k))-h(f(x^*))]=O(\|x_k-x^*\|^2)$ 的方向收敛于 x^* 时,(10.4.14)的误差可导致 $r_k<c_2$, 从而信赖域半径被缩小,因此会出现仅仅线性收敛.

另一种线性收敛的情形和约束优化的马洛托斯效应(Maratos effect)十分类似. 利用例子(10.4.6)～(10.4.8), Yuan (1984b)构造出点列 $\{x_k\}$ 具有下列性质:在第 $2k$ 次迭代,求出的试探步 d_{2k} 是超线性收敛步,即

$$\frac{\|x_{2k}+d_{2k}-x^*\|}{\|x_{2k}-x^*\|}\to 0, \tag{10.4.15}$$

但是 $r_{2k}<0$, 于是有

$$x_{2k+1}=x_{2k}, \tag{10.4.16}$$

$$\Delta_{2k+1}=\frac{1}{2}\|d_{2k}\|_\infty, \tag{10.4.17}$$

在第 $2k+1$ 次迭代有

$$x_{2k+2}=x_{2k+1}+d_{2k+1}=\frac{1}{2}x_{2k}, \tag{10.4.18}$$

$$\Delta_{2k+2}\geqslant\|d_{2k+1}\|_\infty=\Delta_{2k+1}=\|x_{2k+2}\|_\infty. \tag{10.4.19}$$

从而得到一串两步线性收敛的点列 $\{x_k\}$.

§10.5　一个超线性收敛算法

本节讨论一个二阶校正步(second order correction step)的信赖域法. 该方法是由 Fletcher (1982b)提出的. 所以在本节中我们采用 Fletcher (1982b)的记号.

考虑的极小值问题是复合 NDO 问题

$$\min_{x\in \mathbf{R}^n} F(x)=f(x)+h(c(x)), \tag{10.5.1}$$

其中 $f(x)$ 是 $\mathbf{R}^n$ 中的二次连续可微函数, $c(x)$ 是从 $\mathbf{R}^n$ 到 $\mathbf{R}^m$ 的二次连续可微函数, $h(c)$ 是 $\mathbf{R}^m$ 中的凸的分片线性函数 (polyhedral convex function):

$$h(c)=\max_{1\leqslant i\leqslant I}(h_i^T c+b_i), \tag{10.5.2}$$

h_i $(i=1, \cdots, I)$ 和 b_i $(i=1, \cdots, I)$ 分别是给定的向量和常数. 这类特殊形式的复合 NDO 问题(10.5.1)与(10.5.2) 吸引人们的一个主要原因是 L_1 精确罚函数可表示为这种形式.

二阶校正步方法基本上和算法 10.3.1 一样, 只是在一些迭代中多解一个二阶校正子问题而已. 在第 k 次迭代, 假定 $x_k\in \mathbf{R}^n$, $\lambda_k\in \mathbf{R}^m$, $\Delta_k>0$ 已具有, 算法需求解子问题

$$\min_{d\in \mathbf{R}^n} \phi_k(d)=Q_k(d)+h(c_k+A_k d) \tag{10.5.3}$$

$$\text{s. t. } \|d\|\leqslant \Delta_k. \tag{10.5.4}$$

其中

$$Q_k(d)=g_k^T d+\frac{1}{2} d^T w_k d, \tag{10.5.5}$$

$$w_k=\nabla^2 f(x_k)+\sum_{i=1}^{m}(\lambda_k)_i \nabla c_i(x_k) \tag{10.5.6}$$

以及

$$g_k=g(x_k)=\nabla f(x_k),$$

$$c_k=c(x_k),\ A_k=A(x_k)=(\nabla c(x_k)^T)^T.$$

设 d_k 是(10.5.3)与(10.5.4)的解, 且定义

$$r_k=\frac{F(x_k)-F(x_k+d_k)}{\phi_k(0)-\phi_k(d_k)}. \tag{10.5.7}$$

在一些迭代中, 算法需要求解下面的"二阶校正"子问题:

$$\min_{d\in \mathbf{R}^n} \hat{\phi}_k(d)=Q_k(d_k+d)+h(c(x_k+d_k)+A_k d). \tag{10.5.8}$$

$$\text{s. t. } \|d_k+d\|\leqslant \Delta_k. \tag{10.5.9}$$

记 $\hat{d}_k$ 为(10.5.8)与(10.5.9)的解, 定义

$$r_e^{(k)}=r_k+\frac{\hat{\phi}_k(0)-\hat{\phi}_k(\hat{d}_k)}{\phi_k(0)-\phi_k(d_k)},\qquad(10.5.10)$$

$$\hat{r}_k=\frac{F(x_k)-F(x_k+d_k+\hat{d}_k)}{\phi_k(0)-\phi_k(d_k)}.\qquad(10.5.11)$$

利用这些记号，二阶校正步信赖域法可写成下面的形式：

算法 10.5.1

步 1　给出 x_1, λ_0, $\Delta_1>0$, $\varepsilon\geqslant 0$, $k:=1$;

步 2　求解(10.5.3)与(10.5.4)，给出 d_k; 如果 $\|d_k\|\leqslant\varepsilon$，则停;

步 3　计算 r_k，如果 $r_k>0.75$，则转步 8; 解(10.5.8)与(10.5.9)，给出 $\hat{d}_k$; 计算 $r_e^{(k)}$，如果 $r_k<0.25$，则转步 5;

步 4　如果 $x_e^{(k)}\notin[0.9,\ 1.1]$，则转步 9; $\Delta_{k+1}=2\Delta_k$; 转步 10;

步 5　如果 $r_e^{(k)}\notin[0.75,\ 1.25]$，则转步 6; 计算 $\hat{r}_k$; $d_k:=d_k+\hat{d}_k$; $r_k:=\hat{r}_k$; 如果 $r_k>0.75$，则转步 8; 如果 $r_k\geqslant 0.25$，则转步 9;

步 6　令 $\Delta_{k+1}=\alpha_k\|d_k\|$, $\alpha_k\in[0.1,\ 0.5]$; 如果 $r_k>0$，则转步 10;

步 7　$x_{k+1}:=x_k$; 计算 λ_k，转步 11;

步 8　如果 $\|d_k\|<\Delta_k$，则转步 9; 如果 $r_k>0.9$，则 $\Delta_{k+1}=4\Delta_k$，否则，$\Delta_{k+1}=2\Delta_k$; 转步 10;

步 9　$\Delta_{k+1}=\Delta_k$;

步 10　$x_{k+1}=x_k+d_k$; 计算 λ_k;

步 11　$k:=k+1$; 转步 2.

在算法的步 6 中，α_k 是 α_k^* 的一个近似估计，这里 α_k^* 是使得

$$F(x_k+\alpha_k^* d_k)=\min_{0.1\leqslant\alpha\leqslant 0.5}F(x_k+\alpha d_k)\qquad(10.5.12)$$

成立的步长。关于 α_k 的一种具体选取方式可见 Fletcher(1982b)。另外，在步 7 和步 10 中的乘子 λ_k 可取子问题(10.5.3)与(10.5.4)或子问题(10.5.8)与(10.5.9)的 Lagrange 乘子。

关于算法的全局收敛性分析可见 Fletcher (1982b)。我们考虑算法的局部收敛速度，假定由算法产生的点列 x_k 收敛于 x^*，而

且还作如下假定:

假定 10.5.2

1) $f(x)$和 $c_i(x)$为三次连续可微;

2) Rank $(A^*)=m$;

3) $h(c(x^*))=h_i^Tc(x^*)+b_i$ 对所有 $i=1, \cdots, I$ 都成立;

4) 严格互补条件和二阶充分条件在 x^* 处成立.

由于 h 的特殊形式,不难证明

$$\partial h^*=\left\{\lambda=\sum_{i=1}^{I}t_ih_i \,\middle|\, t_i\geqslant 0,\ \sum_{i=1}^{I}t_i=1\right\}. \tag{10.5.13}$$

由假定 10.5.2 的 2)可知道存在唯一的 Lagrange 乘子 λ^*, 使得 $\lambda^*\in\partial h^*$, 且

$$g^*+A^*\lambda^*=0. \tag{10.5.14}$$

我们这里所说的二阶充分条件是指:对任何

$$d\in G^*=\{d,\max_{\lambda\in\partial h^*}d^T[g^*+A^*\lambda]=0, d\neq 0\} \tag{10.5.15}$$

都有

$$d^TW^*d>0, \tag{10.5.16}$$

其中

$$W^*=\nabla^2 f(x^*)+\sum_{i=1}^{m}\lambda_i^*\nabla^2 c_i(x^*). \tag{10.5.17}$$

在以上的假定下,如果对所有充分大的 k 信赖域约束是非积极的,则必有 x_k 二次收敛于 x^*(见 Fletcher,1982b). 但是,从上一节的例子可知,超线性收敛的关键在于信赖域约束是否是积极的. 下面,我们证明只要 Lagrange 乘子的估计较好,则对所有充分大的 k 信赖域约束都不会是积极的,从而可知算法是二次收敛的.

我们假定算法产生的乘子 λ_k 满足

$$\lim_{k\to\infty}\lambda_k=\lambda^*, \tag{10.5.18}$$

估计 Lagrange 乘子的方法很多,如见 Murray 和 Overton(1980, 1981). 另外,简单地令 λ_k 是问题

$$\min\|g_k+A_k\lambda\|_2 \tag{10.5.19}$$

的解也能满足(10.5.18). 所以在下面我们假定(10.5.18)恒成立. 值得指出的是:Fletcher (1982b)提出的选取 λ_k 的方法不一

定满足(10.5.18).

引理 **10.5.3**(Fletcher, 1982b) 存在常数 $\delta_1>0$, 使得在 x^* 为中心的一个小邻域内都有

$$F(x)-F(x^*)\geqslant\delta_1\|x-x^*\|^2. \tag{10.5.20}$$

引理 **10.5.4**(Yuan, 1985b) 对任给 $\alpha_0\in(0, 1)$, 存在一个以 x^* 为中心的小邻域, 使得对该邻域内的 x 均有

$$F(x)-\min_{\|d\|\leq\|x-x^*\|}[F(x)+g(x)^Td+h(c(x)+A(x)^Td)]$$
$$\geqslant\alpha_0[F(x)-F(x^*)]. \tag{10.5.21}$$

引理 **10.5.5**(Yuan, 1985b) 设 d_k 是(10.5.3)与(10.5.4)的解, $\hat{d}_k$ 是(10.5.8)与(10.5.9)的解, 则必有

1) $\|d_k\|=O(\|x_k-x^*\|)$;

2) $\|\hat{d}_k\|=O(\|x_k-x^*\|)$;

3) $\phi_k(0)-\phi_k(d_k)\geqslant\delta_2\|d_k\|^2$, $\delta_2>0$ 是一个常数;

4) $r_e^{(k)}\to 1$;

5) 对任何 $\delta_3<1$ 以及满足 $r_{k_j}\leqslant\delta_3$ 的子列 $\{k_j\}$ 都有

$$\|\hat{d}_{k_j}\|=o(\|d_k\|), \tag{10.5.22}$$

$$\lim_{j\to\infty}\hat{r}_{k_j}=1. \tag{10.5.23}$$

推论 **10.5.6** 对所有充分大的 k 都有 $\Delta_{k+1}\geqslant\Delta_k$.

证明 令引理 10.5.5 中的 $\delta_3=0.75$, 则知对充分大的 k, 如果 $r_k<0.75$, 则有

$$r_e^{(k)}\in[0.9, 1.1], \tag{10.5.24}$$

$$\hat{r}_k\in[0.9, 1.1]. \tag{10.5.25}$$

于是由算法可知 $\Delta_{k+1}\geqslant\Delta_k$. ▌

定理 **10.5.7** 设算法 10.5.1 产生的点列 $\{x_k\}$ 收敛于 x^*, 假定 10.5.2 成立, 以及(10.5.18)成立, 则 x_k 超线性收敛于 x^*.

证明 由推论 10.5.6 可知存在 $\delta_3>0$, 使得

$$\Delta_k\geqslant\delta_3 \tag{10.5.26}$$

对一切 k 都成立. 从引理 10.5.5 的 1) 知 $\|d_k\|\to 0$. 于是对所有充分大的 k, 必有

$$\|d_k\| < \frac{1}{2}\delta_3 < \Delta_k. \tag{10.5.27}$$

所以对所有充分大的 k 信赖域约束都不是积极的. 由 Fletcher (1982b)的结果可证

$$\frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|} = O(\max[\|x_k-x^*\|,\ \|\lambda_k-\lambda^*\|]). \tag{10.5.28}$$

故知 x_k 超线性收敛于 x^*. ∎

推论 10.5.8 如果 λ_k 由(10.5.19)定义, 则 x_k 二次收敛于 x^*.

证明 由于 A^* 是列满秩的, 所以对充分大的 k 均有

$$\lambda_k = -(A_k)^+ g_k.$$

因为 $\lambda^* = -(A^*)^+ g^*$ 和 A^* 列满秩, 我们有

$$\|\lambda_k-\lambda^*\| = O(\|x_k-x^*\|). \tag{10.5.29}$$

由(10.5.28)和(10.5.29)即知 x_k 二次收敛. ∎

稍加修改 Yuan (1985b)的证明, 不难证明在算法中 w_k 用满足

$$B_k \to W^*, \quad 当\ k\to\infty \tag{10.5.30}$$

的矩阵 B_k 代替算法时超线性收敛结果仍然成立. 而且条件(10.5.30)还可放宽到对 $i=1, 2, 3$ 有

$$\|(B_k-W^*)z_i^{(k)}\| / \|z_i^{(k)}\| \to 0. \tag{10.5.31}$$

其中 $z_1^{(k)} = d_k$, $z_2^{(k)} = \hat{d}_k$, $z_3^{(k)} = x_k - x^*$.

在实用算法中, B_k 可由第四章的拟牛顿公式修正产生. 例如 PSB 公式对于非光滑问题具有以下形式:

$$B_{k+1} = B_k - \frac{\eta_k s_k^T + s_k\eta_k^T}{\|s_k\|_2^2} - \frac{\eta_k^T s_k s_k s_k^T}{\|s_k\|_2^4}, \tag{10.5.32}$$

其中

$$s_k = \begin{cases} d_k + \hat{d}_k, & 如果\ x_{k+1} = x_k + d_k + \hat{d}_k; \\ d_k, & 否则. \end{cases} \tag{10.5.33}$$

$$\eta_k = g(x_k+s_k) - g(x_k) - B_k s_k + \sum_{i=1}^{m}(\lambda_k)_i[\nabla c_i(x_k+s_k) - \nabla c_i(x_k)]. \tag{10.5.34}$$

我们建议用 PSB 修正公式的理由之一是将它应用到光滑无约束优化的信赖域法效果很好(见 Powell, 1975).

如果(10.5.4)和(10.5.9)中的范数$\|\cdot\|$取为$\|\cdot\|_\infty$,则(10.5.3)与(10.5.4)和(10.5.8)与(10.5.9)都可化成二次规划问题. 而且这两个二次规划的差别只是线性约束的右端项不同. 所以一般说来,求解出(10.5.3)与(10.5.4)之后,再计算(10.5.8)与(10.5.9)的解所需的计算量不大.

我们证明的超线性收敛结果是基于二阶充分条件以及$\nabla c_i(x^*)$ $(i=1, \cdots, m)$线性无关. 显而易见,二阶充分条件是证明超线性收敛结果必不可少的. 作者还认为$\nabla c_i(x^*)$ $(i=1, \cdots, m)$线性无关这一条件也不能放宽.

参考文献

[1] J. Abadie and J. Carpentier, "Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints", in: R. Fletcher, ed., *Optimization* (Academic Press, London, 1969)

[2] S. Agmon, "The relaxation method for linear inequalities", *Canadian J. Math.* 6 (1954) 382~392.

[3] H. Akaike, "On a succesive transformation of probability distribution and its application to the analysis to. the optimum gradient method", *Annual Inst. Statist. Math. Tokyo* 11 (1959) 1~16.

[4] M. Al-Baali, "Descent property and global convergence of the Fletcher-Reeves method with inexact line search" *IMA J. Numer. Anal.* 5 (1985) 121~124.

[5] R. C Archibald, "Golden section", *American Math. Monthly* 25 (1918) 232~238.

[6] L. Armijo, "Minimization of functions having Lipschitz continuous partial derivatives", *Pacific J. Math.* 16 (1966) 1~3.

[7] M. L Balinski and P. Wolfe, (eds). *Mathematical Programming Study 3: Nondifferentiable Optimization* (North-Holland, Amsterdam, 1975)

[8] M. S. Bazara and C. M. Shetty, *Nonlinear Programming, Theory and Algorithms* (John Wiley and Sons, New York, 1979)

[9] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods", *IMA J. Numer. Anal.* 8 (1988) 141~148.

[10] E. M. L. Beale, "On quadratic programming", *Naval Res Logistics Quarterly* 6 (1959) 227~243.

[11] C. S. Beighter, D. T. Phillips and D. J. Wilde, *Foundations of Optimization* (Prentice-Hall, Englewood Cliffs, N. J 1979)

[12] E. T. Bell, *The Development of Mathematics* (Mc Graw-Hill, New York, 1940)

[13] D. P. Bertsekas, "Variable metric methods for constrained optimization based on differentiable exact penalty functions", *Proc. Conf. Comm. Control Comput., Allerton Park, III* pp. 584~593, 1980.

[14] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, New York, 1982) (1982a)

[15] D. P. Bertsekas, "Augmented Lagrangian and differentiable exact penalty methods", in: M. J. D. Powell, ed., *Nonlinear Optimization* 1981 (Academic Press, London, 1982) pp. 223~234. (1982b)

[16] P. T. Boggs and J. W. Tolle, "Merit function for nonlinear programming problems", Operations Res. and Sys. Anal. Report No 81-2, University of North Carolina at Capel Hill, 1981.

[17] P. T. Boggs, J. W. Tolle and P. Wang, "On the local convergence of quasi-Newton methods for constrained optimization", *SIAM J. Control Opt.* 20 (1982) 161～171.

[18] P. T. Boggs and J. W. Wang, "An efficient strategy for utilizing a merit function in nonlinear programming algorithms", Report UNC/ORSA/TR-85/5, University of North Carolina, 1985.

[19] G. C. Broyden, "Quasi-Newton methods and their application to function minimization", *Maths. Comput* 21 (1967) 368～381.

[20] G. C. Broyden, "The covergence of a class of double rank minimization algorithms: 2. the new algorithm", *J. Inst. Math. Appl.* 6 (1970) 222～231.

[21] G. C. Broyden, J. E. Dennis and J. J. Moré, "On the local and superlinear convergence of quasi-Newton methods", *J. Inst. Math. Appl.* 12 (1973) 223～245.

[22] A. Buckley, "A combined conjugate gradient quasi-Newton minimization algorithms", *Math. Prog.* 15 (1978) 200～210.

[23] A. Buckley, "Conjugate gradient methods", in: M. J. D. Powell, ed., *Nonlinear Optimisation* 1981 (Academic Press, London, 1982) pp. 17～22.

[24] A. Buckley, A. LeNir, "QN-like variable storage conjugate gradients", *Math. Prog.* 27 (1983) 155～175.

[25] J. Burke, "Descent methods for composite nondifferential optimization problems", *Math. Prog.* 33 (1985) 260～279.

[26] J. Burke, "Second order necessary and sufficient conditions for convex composite NDO", *Math. Prog.* 38 (1987) 287～302.

[27] J. V. Burke, J. J. Moré and G. Toraldo, "Convergence properties of trust region methods for linear and convex constraints", *Math. Prog.* 47 (1990) 305～336.

[28] W. Burmeister, "Die Konvergenzornung des Fletcher-Powell Algorithms", *Z. Angew. Math. Mech.* 53 (1973) 693～699.

[29] R. H. Byrd, D. C. Liu, and R. H. Nocedal, "On the behaviour of Broyden's class of quasi-Newton methods", Technicol Report NAM 01, Department of Electrical Engineering and Computer Science, North-western University, 1990.

[30] R. Byrd and J. Nocedal, "A tool for the analysis of quasi-Newton methods with application to unconstrained minimization", *SIAM J. Numer. Anal.* 26 (1989) 727～739.

[31] R. Byrd and J. Nocedal, "An analysis of reduced Hessian methods for constrained optimization" *Math. Prog*. 49 (1991) 285～323.

[32] R. Byrd, J. Nocedal and Y. Yuan, "Global convergence of a class of variable metric algorithms", *SIAM J. Numer. Anal*. 4 (1987) 1171～1190.

[33] R. Byrd, R. B. Schnabel and G. A. Shultz, "A trust region algorithm for nonlinearly constrained optimization", *SIAM J. Numer. Anal*. 24 (1987) 1152～1170.

[34] P. H. Calamai and Moré, "Projected gradient methods for linearly constrained problems", *Math. Prog*. 39 (1987) 93～116.

[35] C. W. Carroll, "The created responds surface technique for optimizing nonlinear, restrained systems", *Operations Res*. 9 (1961) 169～184.

[36] A. Cauchy, "Méthode générale pour la résolution des systèms d'équations simulatnées", *Comptes Rendus de L'Acadmia Des Sciences* 25 (1847) 536～538.

[37] M. R. Celis, J. E. Dennis and R. A. Tapia, "A trust region algorithm for nonlinear equality constrained optimization", in: P. T. Boggs, R. H. Byrd and R. B. Schnabel, eds., *Numerical Optimization* (SIAM, Philadelphia, 1985) pp. 71～82.

[38] R. M. Chamberlain, "Some examples of cycling in variable metric methods for constrained minimization", *Math. Prog*. 16 (1979) 378～383

[39] R. M. Chamberlain, C. Lemarechal, H. C. Pedersen and M. J. D. Powell, "The watchdog techniques for forcing convergence in algorithms for constrained optimization", *Math. Prog. Study* 16 (1982) 1～17.

[40] C. Charalambous and J. W. Bandler, "Nonlinear minimax optimization as a sequence of least p-th optimization with finite value of p", *Internat. J. Systems Sci*. 7 (1976) 377～391.

[41] C. Charalambous and A. R. Conn, "Optimization of microwave networks" *IEEE Trans. Microwave Theory and Tech*. (1975) 834～838.

[42] C. Charalambous and A. R. Conn. "An efficient method to solve the minimax problem directly", *SIAM J. Numer. Anal*. 15 (1978) 162～187.

[43] A. Charnes, "Optimality and degeneracy in linear programming", *Econometrica* 20 (1952) 160～170.

[44] E. W. Cheney and A. A. Goldstein, "Newton's method for convex programming and Chebyshev approximation", *Numerische Mathematik* 1 (1959) 253～268.

[45] F. H. Clark, "Generalized gradients and applications", *Trans. of Amer. Math. Soc*. 205 (1975) 247～262.

[46] F. H. Clark, *Opitmization and Nonsmooth Analysis*, (John Wiley and

Sons, New York, 1983).

[47] T. Coleman and A. R. Conn, "On the local convergence of a quasi-Newton method for the nonlinear programming problem", *SIAM J. Numer. Anal.* 21 (1984) 755~769.

[48] R. W. Cottle, J. Kyparisis and J. -S. Pang, eds., *Complementarity Problems: Mathematical Programming Series B* 48 (1990) No. 3.

[49] R. Courant, Variational methods for the solution of problems of equilibrium and vibrations", *Bulle tin of American Mathematics Soceity* 49 (1943 1~3.

[50] H. S. M. Coxeter, "The Golden section, phyllotaxis, and Wythoff's game", *Scipta Mathematica* (1954) 135~143.

[51] G. B. Dantzig, A. Orden and P. Wolfe. "The generalized simplex method for minimizing a linear form under linear inequality restraints", *Pacific J. Maths.* 5 (1955) 183~195.

[52] W. C. Davidon, "Variable metric method for minimization", AEC Res. and Dev. Report ANL-5900, 1959. / bibitemdavidon 68 W. C. Davidon, "Variance algorithms for minimization", *Computer J.* 10 (1968) 406~410.

[53] W. C. Davidon, "Optimally conditioned optimization algorithms without line searches" *Math. Prog.* 9 (1975) 1~30.

[54] W. C. Davidon, R. B. Mifflin and J. L. Nazareth, "Some comments on notation for quasi-Newton methods", *Mathematical Programming Society Newsletter*, *OPTIMA* No. 32 (1991) 3~4.

[55] V. F. Demyanov and L. C. W. Dixon, (eds.) *Mathematical Programming Study 29: Quasidifferential Calculus* (North-Hollad, Amsterdam, 1986)

[56] J. B. Dennis, *Mathematical Programming and Electrical Networks* (Technology Press, Cambridge, Mass., 1959)

[57] J. E. Dennis and J. J. Moré, "A chareaterization of superlinear convergence and its application to quasi-Newton methods", *Mathematics of Computation* 28 (1974) 549~560.

[58] J. E. Dennis and J. J. Moré, "Quasi-Newton method, motivation and theory", *SIAM Review* 19 (1977) 46~89.

[59] J. E. Dennis Jr. and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations* (Prentice-Hall, Englewood Cliffs, NJ, 1983)

[60] J. E. Dennis and K. Turner, "Generalized conjugate directions", Report 85-11, Dept of Mathmetics, Rice University, Houston, 1985.

[61] G. Di Pillo and L. Grippo, "A new class of augmented Lagrangians in nonlinear programming", *SIAM J. Control Opt.* 17 (1979) 618~828.

[62] G. Di Pillo and L. Grippo, "An exact penalty function method with global convergence properties for nonlinear programming problem", *Math. Prog.* 36 (1986) 1~18.

[63] L. C. W. Dixon, "The choice of step length, a crucial factor in the performance of variable metric method", in: F. A. Lootsma, ed., *Numerical Methods for Nonlinear Optimization* (Academic Press, London, 1972) pp. 149~170. (1972a)

[64] L. C. W. Dixon, "Variable metric algorithms: necessary and sufficient conditions for identical behavior of nonquadratical functions", *J. Opt. Theory Appl.* 10 (1972) 34~40. (1972b)

[65] L. C. W. Dixon, E. Spedicato and G. P. Szego, (eds.) *Nonlinear Optimization* (Birkhäuser, Boston, 1980)

[66] W. S. Dorn, "Duality in quadratic programming", *Quart. Appl. Math.* 18 (1960) 155~162.

[67] I. I. Eremin, "A generalization of the Motzkin–Agmon relaxation method"., *Soviet Math. Doklady* 6 (1965) 219~221.

[68] Yu. M. Ermoliev, "Method of solution of nonlinear extremal problems" (in Russian), *Kibernetika* 2 (1966) 1~17.

[69] J. Farkas, "Ubar die Theorie der Einfachen Ungleichungen", *J. für die Rein und Angew. Math.* 124 (1902) 1~27.

[70] A. V. Fiacco and G. P. Mc Cormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (John Wiley, New York 1968).

[71] Fibonacci, Leonardo, *Algebra et almuchabala (liber abbaci)*, 1202.

[72] R. Fletcher, "A new approach to variable metric algorithms", *Computer J.* 13 (1970) 317~322.

[73] R. Fletcher, "A general quadratic programming algorithm", *J. Inst. Maths. Applns.* 7 (1971) 76~91.

[74] R. Fletcher, "An exact penalty function for nonlinear programming with inequalities", *Math. Prog.* 5 (1973) 129~150.

[75] R. Flecther, "An ideal penalty function for constrained optimization", *J. Inst. Math. Applns.* 15 (1975) 319~342.

[76] R. Fletcher, *Practical Methods of Optimization, Vol. 1, Unconstrained Optimization* (John Wiley and Sons, Chichester, 1980)

[77] R. Fletcher, *Practical Methods of Optimization, Vol. 2, Constrained Optimization* (John Wiley and Sons, Chichester, 1981)

[78] R. Fletcher, "A model algorithm for composite NDO problem", *Math. Prog. Study* 17 (1982) 67~76. (1982a)

[79] R. Fletcher, "Second order correction for nondifferentiable optimization", in: G. A. Watson, ed., *Numerical Analysis* (Springer–Verlag, Berlin,

1982) pp. 85~115. (1982b)

[80] R. Fletcher, *Practical Methods of Optimization* (second edition) (John Wiley and Sons, Chichester, 1987)

[81] R. Fletcher, "A new variational result for quasi-Newton formulae" Technical Report NA/119, Department of Mathematical Sciences, University of Dundee, 1989.

[82] R. Fletcher and M. J. D. Powell, "A rapid convergent descent method for minimization", *The Computer J.* 6 (1963) 163~168.

[83] R. Fle tcher and C. M. Reeves, "Function minimization by conjugate gradients", *Computer J.* 7 (1964) 149~154.

[84] R. Fletcher and G. A. Watson, "First and second order conditions for a class of nondifferentiable optimization problems", *Math. Prog.* 18 (1980) 291~307.

[85] G. E. Forsythe, "On the asymptotic directions of the s-dimentional optimum gradient method", *Numerische Mathematik* 11 (1968) 57~76.

[86] K. R. Frisch, "The logarithmic potential method of convex programming", Memoradum, University Institute of Economics, Oslo, May, 1959.

[87] M Fukushima, "A succesive quadratic programming algorithm with global and superlinear convergence properties", *Math. Prog* 35 (1986) 253 ~264.

[88] D. M. Gay, "Computing optimal local constrained step", *SIAM J. Sci Stat. Comp.* 2 (1981) 186~197.

[89] E. M. Gafni and D. P. Bertsekas, "Two-metric projection method for constrained optimization", *SIAM J. Control Opt.* 22 (1984) 936~964.

[90] P. E. Gill and W. Murray, "Quasi-Newton methods for unconstrained optimization", *J. Inst. Maths. Applns.* 9 (1972) 91~108.

[91] P. E. Gill and W. Murray, "Conjugate gradient methods for large-scale nonlinear optimization", Technical Report SOL 79-15, Department of Operations Research, Stanford University (Stanford, California, 1979).

[92] P. E. Gill, W. Murray, M. A. Sauders and M. H. Wright, "User's guide for SOL/NPSOL: A FORTRAN package for nonlinear programming", Techical Report SoL 83~12, Dept of OR, Stanford University, Stanford, 1983

[93] J. L. Goffin, "Nondifferentiable optimization and the relaxation method", in: C. Lemaréchal and R. Mifflin, eds., *Nonsmooth Optimization* (Pergamon, Oxford, 1978) pp. 31~49.

[94] D. Goldfarb, "A family of variable metric method derived by variation mean", *Math. Comput.* 23 (1970) 23~26

[95] D. Goldfarb and A. Idinani, "A numerical stable dual method for

solving strictly convex quadratic programs", *Math. Prog.* 27(1983) 1～33.

[96] D. Goldfarb and S. Liu, "An $O(n^3L)$ primal interior point algorithm for convex quadratic programming", *Math. Prog.* 49(1991) 325～340.

[97] A. A. Goldstein, "On steepest descent", *SIAM J. Control* 3(1965) 147～151.

[98] A. A. Goldstein and J. F. Price, "An effective algorithm for minimization", *Numerische Mathematik* 10(1967) 184～189.

[99] J. Greenstadt, "A quasi-Newton method with no derivatives" *Mathematics of Computaions* 26(1972) 145～166. (1972a)

[100] J. Greenstadt, "Improvements in a QNWD meth od" IBM Palo Alto Scientific Center Technical Report No. 320-33, Palo Alto, California, 1972. (1972b)

[101] M. Guignard, ed., *Mathmatical Programming Study* 19: *Optimality and Stability in Mathematical Programming* (North-Holland, Amsterdam, 1982)

[102] J. Hald and K. Madsen, "A 2-stage algorithm for minimax optimization", in: *International Symposium on Systems Optimization and Analysis* (Springer-Verlag, Berlin, 1979) pp. 225～239.

[103] J. Hald and K. Madsen, "Combined LP and quasi-Newton methods for minimax", *Math. Prog.* 20(1981) 49～62.

[104] S. P. Han, "A global convergent method for nonlinear programming", *J. Opt. Theory Appl.* 22(1977) 297～309.

[105] S. P. Han and O. L Mangasarian, "A dual differentiable exact penalty function", *Math. Prog.* 25(1983) 293～306.

[106] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford University Press, Oxford, 1971)

[107] J. Helmes, "Eine einfachere, auf einer neuen Analyse beruhende Auflösung der sectio aurea, nebst einer kritischen Beleuchtung der gewöhnlichen Auflösung und der Betrachtung ihres pädagogischen Werthes", *Archiv der Mathematik* 4(1844) 15～22.

[108] M. R. Hestenes, "Multiplier and gradient methods", *J. Opt. Theory Appl.* 4(1969) 303～320.

[109] M. R. Hestenes and E. Stiefel, "Method of conjugate gradient for solving linear system", *J. Res. Nat. Bur. Stand.* 49(1952) 409～436.

[110] R. Hooke and T. A. Jeeves, "Direct search solution of numerical and statistical problems", *J. ACM* 8(1961) 212～229.

[111] R. Hornung, "Discret minimax problems: algorithms and numerical comarisons", *Computing* 28 (1982) 139～154. (1982a)

[112] R. Hornung, "Ein 2-stufen Newton -Demyanov Algorithmus zur Losung eines diskreten Minimaxproblemes", *Math. Operationsforsch Statist. Ser. Optimization* 13(1982) 77～92. (1982b)

[113] R. Hornung, "An efficient algorithm for solving the discret minisum problem", *Operations Research Letters* 2(1983) 115～118.

[114] S. Hoshino, "A formulation of variable metric methods", *J. Inst. Maths. Applns.* 10(1972) 394～403.

[115] H. Y. Huang, "Unified approach to quadratically convergent algorithms for function minimization", *J. Opt. Theory and Appl.* 5(1970) 405～423.

[116] S. Kapoor and P. Vaidya, "Fast algorithms for convex quadratic programming and multicommodity flows", *Proceeding of the 18th Annual ACM Symp. Theory Comput.*, (1986) pp. 147～159.

[117] N. Karmarkar, "A new polynomial-time algorithm for linear programming", *Combinatorica*, 4(1984) 374～395.

[118] W. Karush, *Minima of functions of several variables with inequalities as side conditions*, Master's thesis, University of Chicago, Chicago, Illinois, 1939.

[119] J. E. Kelley, "The cutting plane method for solving convex programs", *J. of SIAM* 8(1960) 703～712.

[120] H. W. Kuhn and A. W. Tucker, "Nonlinear programming", in: J. Neyman, ed., *Proceeding of the Second Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, California, 1951) pp. 481～492.

[121] C. J. L. Lagrange, "Essai dune nouvelle methods pour deteminer les maxima et les minima", *Miscellanea Taurinensia* 2(1760-61) *Oeuvres*, 1, pp. 356～57, 360.

[122] M. Lalee and J. Nocedal, "Automatic column scaling strategies for quasi-Newton methods", Report NAM 04, EECS Dept, Northwestern University, USA, 1991.

[123] C. Lemaréchal, "Bundle methods in nonsmooth optimization", in: C Lemaréchal and R. Mifflin, eds., *Nonsmooth Optimization* (Pergamon, Oxford, 1978) pp. 79～102. (1978a)

[124] C. Lemaréchal, "Nonsmooth optimization and descent methods", RR-78-4, IIASA report, 1978. (1978b)

[125] C. Lemaréchal, "Nondifferentiable optimization", in: L. C. W. Dixon, E. Spedicato and G. P. Szego, eds., *Nonlinear Optimization* (Birkhäuser, Boston, 1980) pp. 149～199.

[126] C. Lemaréchal and R. Mifflin, (eds.) *Nonsmooth Optimization* (Perga-

mon, Oxford, 1978).

[127] C. Lemaréchal, J. J. Strodiot and A. Bibain, "On a bundle algorithm for nonsmooth optimization", in: O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds., *Nonlinear Programming* 4 (Academic Press, New York, 1981) pp. 245～282.

[128] C. E. Lemke, "A method of solution of quadratic programs", *Management Sci*. 8 (1962) 442～455.

[129] C. E. Lemke, "Bimatrix equilibrium points and mathematical programming", *Management Sci*. 11 (1965) 681～689.

[130] F. Lempio and J. Zowe, "Higher order optimality conditions", in: B. Korte, ed, *Operations Research* (North-Holland, Amsterdam, 1982).

[131] K. Levenberg, "A method for the solution of certain nonlinear problems in least squares", *Qart. Appl. Math*. 2 (1944) 164～166.

[132] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization", *Math. Prog*. 45 (1989) 503～528.

[133] D. G. Luenberger, *Linear and Nonlinear Programming* (2nd Edition) (Addison-Wesley, Massachusetts, 1984).

[134] C. P. Mc Cormick, *Nonlinear Programming: Theory, Algorithms, and Applications*. (John Wiley and Sons, New York, 1983)

[135] K. Madsen, "An algorithm for the minimax solution of overdetermined systems of nonlinear equations", *J. Inst. Math. Appl*. 16 (1975) 1～20.

[136] O. L. Mangasarian and S. Fromowitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints", *J. Math. Anal. Appl*. 17 (1967) 37～47.

[137] N. Maratos, *Exact Penalty Function Algorithms for Finite Dimensional and Control Optimization Problems*, Ph. D. thesis, Imperial College Sci. Tech., University of London, 1978.

[138] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear inequalities", *SIAM J. Appl. Math*. 11 (1963) 431～441.

[139] J. M. Martinez, "Local convergence theory of inexact Newton methods based on structed least change updates", *Mathematics of Computation* 55 (1990) 143～167.

[140] D. Q. Mayne and E. Polak, "A superlinearly convergent algorithm for constrained optimization problems", *Math. Prog. Study* 16 (1982) 45～61.

[141] S. Mehrotra and J. Sun, "An algorithm for convex quadratic programming that requires $O(n^{3.5}L)$ arithmetic operations" *Math. of Oper. Res*. 15 (1990) 342～363.

[142] R. C. Monteiro and I. Adler, "Interior path following primal-dual

algorithms. Part II: convex quadratic programming", *Math. Prog.* 44 (1989) 42~66.

[143] J. J. Moré, "Recent developments in algorithms and software for trust region methods", in: A. Bachem, M. Grötschel and B. Korte, eds., *Mathematical Programming The State of the Art* (Springer-Verlag, Berlin, 1983) pp. 258~287.

[144] J. J. Moré and D. C. Sorenson, "Computing a trust region step", *SIAM J. Sci. Stat. Comp.* 4(1983) 553~572.

[145] T. Motzkin and I. J. Schoenberg, "The relaxation method for linear inequalities", *Candian J. Math.* 6(1954) 393~404.

[146] W. Murray, "An algorithm for finding a local minimum of an indefinite quadratic program", NPL report NAC 1, 1971.

[147] W. Murray and M. L. Overton, "A projected Lagrangian algorithm for nonlinear minimax optimization", *SIAM J. Sci. Stat. Comp.* 1 (1980) 345~370.

[148] W. Murray and M. L. Overton, "A projected Lagrangian algorithm for nonlinear L_1 optimization", *SIAM J. Sci. Stat. Comp.* 2(1981) 207 ~224.

[149] B. A. Murtagh and R. W. H. Sargent, "A constrained optimization method with quadratic convergence", in: R. Fletcher, ed., *Optimization* (Academic Press, London, 1969) pp. 215~346.

[150] G. E. Myers, "Properties of conjugate gradient and Davidon methods", *J. Opt. Theory Appl.* 2(1968) 209~219.

[151] J. A. Nelder and R. Mead, "A simplex method for function minimization", *The Computer J.* 7(1965) 308~313.

[152] J. Nocedál, "Updating quasi-Newton matrices with limited storage", *Mathematics of Computation* 35(1980) 773~782.

[153] J. Nocedal and M. L. Overton, "Projected Hessian update algorithms for nonlinear constrained optimization", *SIAM J. Numer. Anal.* 22 (1985) 821~850.

[154] E. A. Nurminski, (ed.) *Progress on Nondifferentiable Optimization* (IIASA, Laxenburg, 1982)

[155] S. S. Oren, *Self-Scaling Variable Metric Algorithms for Unconstrained Minimization* Ph. D. Thesis, Department of Engineering Economics Systems, Stanford University, 1972.

[156] S. S. Oren, "Perspectives on self-scaling variable metric algorithms", *J. Opt. Theory Appl.* 37(1982) 137~147.

[157] S. S. Oren and D. G. Luenberger, "Self-scaling variable metric (SSVM) algorithms, part I Criteria and sufficient conditions for scaling

a class of algorithms", *Manage. Sci.* 20 (1974) 845～862.

[158] J. R. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, London, 1970)

[159] M. R. Osborne and G. A. Watson, "An algorithm for minimax approximation in nonlinear case", *The Computer J.* 12 (1969) 63～68.

[160] M. L. Overton, "Algorithms for nonlinear l_1 and l_∞ fitting", in: M. J. D. Powell, ed., *Nonlinear Optimization 1981* (Academic Press, London, 1982).

[161] J. R. Palmer, "An improved procedure for orthogonalising the search vectors on Rosenbrock's and Swann's direct search optimization methods", *Computer J.* 12 (1969) 69～71.

[162] J. D. Pearson, "Variable metric methods of minimization", *The Computer J.* 12 (1969) 171～178.

[163] J. M. Perry, "A class of conjugate gradient algorithms with a two-step vaiable-metric memory", Disscussion Paper 269, Center for Mathematical Studies in Economics and Management Science, Northwestern University (Evanston, Illinois, 1977).

[164] E. Polak and G. Ribiére, "Note sur la convergence de méthodes de directions conjugees" *Rev. Fr. Inform. Rech. Oper.* 16 (1969) 35～43.

[165] B. T. Polyak, "A general method of solving extremal problems", *Soviet Math. Doklady* 8 (1967) 14～29.

[166] B. T. Polyak, "The conjugate gradient method in extremum problems", *USSR Comp. Math. and Math. Phys.* 9 (1969) 94～112.

[167] B. T. Polyak, "Subgradient methods: A survey of Soviet research", in: C. Lemaréchal and R. Mifflin, eds., *Nonsmooth Optimization* (Pergamon, Oxford, 1978) pp. 5～30.

[168] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives", *The Computer J.* 7 (1964) 155～162.

[169] M. J. D. Powell, "On the calculation of orthogonal vectors", *Computer J.* 11 (1968) 302～304.

[170] M. J. D. Powell, "A theory on rank one modifications to a matrix and its inverse", *The Computer J.* 12 (1969) 288～290.

[171] M. J. D. Powell, "A new algorithm for unconstrained optimization", in: J. B. Rosen, O. L. Mangasarian and K. Ritter, eds., *Nonlinear Programming* (Academic Press, New York, 1970) pp. 31～66.

[172] M. J. D. Powell, "On the convergence of the variable metric algorithm", *J. Inst. Maths. Appl.* 7 (1971) 21～36.

[173] M. J. D. Powell, "Quadratic termination properties of minimization

algorithms, Part I and Part II", *J. Inst. Maths. Appl.* 10 (1972) 332～357.

[174] M. J. D. Powell, "Convergence properties of a class of minimization algorithms", in: O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds., *Nonlinear Programming* 2 (Acadmic Press, New York, 1975) pp. 1～27.

[175] M. J. D. Powell, "Some global convergence properties of a variable metric algorithm for minimization without exact line searches", in: R. W. Cottle and C. E. Lemke, eds., *Noninear Programming, SIAM-AMS Proceegins vol. IX* (SIAM publications, Philadelphia, 1976) pp. 53～72. (1976a)

[176] M. J. D. Powell, "Some convergence properties of the conjugate gradient method", *Math. Prog.* 11 (1976) 42～49.

[177] M. J. D. Powell, "Restart procedure for the conjugate gradient method" *Math. Prog.* 12 (1977) 241～254.

[178] M. J. D. Powell, "A fast algorithm for nonlinearly constrained optimization calculations", in: G. A. Watson, ed., *Numerical Analysis* (Springer-Verlag, Berlin, 1978) pp. 144～157.

[179] M. J. D. Powell, "Nonconvex minimization calculations and the conjugate gradient method", in: D. F. Griffiths, ed., *Numerical Analysis* Lecture Notes in Mathematics 1066 (Springer-Verlag, Berlin, 1984) pp. 122～141. (1984a)

[180] M. J. D. Powell, "On the rate of convergence of variable metric algorithms for uncosntrained optimization", in: Z. Ciesielki and C. Olech, eds., *Proceeding of the International Congress of Mathematicians* (Elsevier, New York, 1984) pp. 1525～1539. (1984b)

[181] M. J. D. Powell, "General algorithms for discrete nonlinear approximation calculations", in: L. L. Schumarker, ed., *Approximation Theory IV* (Academic Press, New York, 1984) pp. 187～218. (1984c)

[182] M. J. D. Powell, "On the global convergence of trust region algorithms for unconstrained optimization", *Math. Prog.* 29 (1984) 297～303. (1984d)

[183] M. J. D. Powell, "On the quadratic programming algorithm of Goldfarb and Indani", *Math. Prog. Study* 25 (1985) 46～61.

[184] M. J. D. Powell, "How bad are the bfgs and dfp methods when the objective function is quadratic", *Math. Prog.* 34 (1986) 34～47.

[185] M. J. D. Powell, "Updating conjugate directions by the BFGS formular", Math. Prog. 38 (1987) 29～46.

[186] M. J. D. Powell, "A tolerant algorithm for linearly constrained optimi-

zation calculations", *Math. Prog.* 45(1989) 547～566.

[187] M. J. D. Powell and Y. Yuan, "Conditions for superlinear convergence in l_1 and l_∞ solutions of overdetermined nonlinear equations", *IMA J. Numer. Anal.* 4(1984) 241～251.

[188] M. J. D. Powell and Y. Yuan, "A recursive quadratic programming algorithm that use differentiable exact penalty function", *Math. Prog.* 35(1986) 265～278.

[189] M. J. D. Powell and Y. Yuan, "A trust region algorithm for equality constrained optimization", *Math. Prog.* 49(1991) 189～211.

[190] K. Ritter, "On the rate of superlinear convergence of a class of variable metric methods", *Numer. Math.* 35(1980) 293～313.

[191] R. T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970).

[192] R. T. Rockafellar, "The multiplier method of Hestenes and Powell applied to convex programming", *J. Opt. Theory Appl.* 12(1973) 555～562.

[193] R. T. Rockafellar, *The Theory of Subgradient and Its Application to Problems of Optimization: Convex and Not Convex Functions* (Hèldermann Verlag, West Berlin, 1981).

[194] J. B. Rosen, "The gradient projection method for nonlinear programming, Part 1: Linear constraints", *J. SIAM* 8(1960) 181～217.

[195] H. H. Rosenbrock, "An automatic method for finding the greatest or least value of a function", *Computer J* 3(1960) 175～184.

[196] K. Schittkowski, "The nonlinear programming method of Wilson, Han and Powell with an augmented Lagragian type line search function, Part 1: convergence analysis", *Numer. Math.* 38(1981) 83～114.

[197] K. Schittkowski, "On the convergence of a sequential quadratic programming method with an augmented Lagragian line search function", *Math. Operationsforsch. u. Statist., Ser Optimization* 14(1983) 197～216.

[198] G. Schuller, "On the order of convergence of certain quasi-Newton methods", *Numer. Math.* 23(1974) 181～192.

[199] G. Schuller and J. Stoer, "Über die Konvergenzordnung gewisser Rang 2 Verfahren zur Minimierung von Funktionen", in: *International series of Numerical Mathematics*, Vol 23 (Birkhäuser, Basel, 1974) pp. 125～147.

[200] D. F. Shanno, "Conditioning of quasi-Newton methods for function minimization", *Math. Comput.* 24(1970) 647～656.

[201] D. F. Shanno, "Conjugate gradient methods with inexact searches",

Math. Oper. Res. 3(1978) 244~256.

[202] D. F. Shanno and K. H. Phua, "Matrix conditioning and nonlinear optimization", *Math. Prog.* 14(1978) 149~160.

[203] N. Z. Shor, "Application of the gradient method for solution of network transportation problems" (in Russian), in: *Notes Scientific Seminar on Theory and Application of Cybernetics and Operations Research*, (Acad. of Sciences, 1962).

[204] N. Z. Shor, "The rate of convergence of the generalized gradient descent method" (in Russian), *Kibernetika* 4(1968) 98~99.

[205] N. Z. Shor, "Utilization fo the operation of space dilatation in the minimization of convex function" (in Russian), *Kibernetika* 6(1970) 6~12.

[206] N. Z. Shor, "Cut-off method with space extension in convex programming problems" (in Russian), *Kibernetika* 13(1977) 94~95.

[207] N. Z. Shor, "Generalized gradient methods of nondifferentiable optimization emplying space dilatation operations", in: A Bachem. M. Grötschel and B. Korte, eds., *Mathematical Programming: The State of the Art* (Springer-Verlag, Berlin, 1983) pp. 501~529.

[208] V. A. Skokov. "Note on minimization methods using operation of space dilatation" (in Russian), *Kibernetika* 10(1974) 115~117.

[209] . . S. Smith, "The automatic computation of maximum likehood estimates", NCB Sci. Dept. Report SC 864/MR/40, 1960.

[210] G. Sonnevend, "An analytic center' for polybedrons and new classes of global algorithms for linear (smooth, convex) programming", *Proceedings of the 12th IFIP Conference on System Modeling and Optimization* (Budapest, 1985).

[211] D. C. Sorensen and R. J. B. Wets, (eds.) *Mathematical Programming Study 17: Nondifferetial and Variational Techniques in Optimization* (North-Holland, Amsterdam, 1982)

[212] W. Spendley, G. R. Hext and F. R. Himsworth, "Sequential application of simplex designs in optimization and Evolutionary operation", *Technometrics* 4(1962) 441~461.

[213] G. W. Stewart, "A modification of Davidon's minimization method to accept difference approximation of derivativos", *J. Ass. Comput. Math.* 14(1967) 72~83.

[214] W. H. Swann, "Report on the development of a new direct serach method of optimization", ICI L td. Cent. Inst. Lab. Research Note 64/3 (1964).

[215] W. H. Swann, "Direct search methods", in: W. Murray, ed. *Numerical*

Methods for Unconstrained Optimization (Academic Press, London, 1972) pp. 13～28.

[216] J. J. Sylvester, "A question in the geometry of situation", *Quart. J. Math.* 1 (1857) 79.

[217] P. Wolfe, "Methods of Recent advances in mathematical programming", in: R. L. Graves and P. Wolfe, eds., *Recent Advances in Mathematical Programming* (McGraw-Hill, New York, 1963) pp. 67～86.

[218] P. Wolfe, "Another variable metric method", working paper, 1968.

[219] P. Wolfe, "A method of conjugate subgradients for minimizing nondifferentiable functions", *Math. Prog. Study* 3 (1975) 145～173.

[220] R. S. Womersley, "Local properties of algorithms for minimizing nonsmooth composite functions", *Math. Prog.* 32 (1985) 69～89.

[221] G. A. Watson, "A class of programming problems whose objective function contains a norm", *J. Approximation Theory* 23 (1978) 401～411.

[222] G. A. Watson, "The minimax solution of an overdetermined systems of nonlinear equations", *J. Inst. Math. Appl.* 23 (1979) 167～180.

[223] Y. Ye and M. J. Todd, "Containing and shrinking ellipsoids in the path-following algorithm", *Math. Prog.* 47 (1990) 1～9.

[224] Y. Ye and E. Tse, "An extention of Karmarkar's algorithm to convex quadratic programming" *Math. Prog.* 44 (1989) 157～179.

[225] Y. Yuan, "Some properties of trust region algorithms for nonsmooth optimization", Report DAMTP 1983/NA4, University of Cambridge, England. (1983a)

[226] Y. Yuan, "Global convergence of trust region algorithms for nonsmooth optimization", Report DAMTP 1983/NA13, University of Cambridge, England. (1983b)

[227] Y. Yuan, "On the least Q-order of convergence of variable metric algorithms", *IMA J. Numer. Anal.* 4 (1984) 233～239. (1984a)

[228] Y. Yuan, "An example of only linearly convergence of trust region algorithms for nonsmooth optimization", *IMA J. Numer, Anal.* 4 (1984) 327～335. (1984b)

[229] Y. Yuan, "Conditions for convergence of trust region algorithms for nonsmooth optimization", *Math. Prog.* 31 (1985) 220～228. (1985a)

[230] Y. Yuan, "On the superlinear convergence of a trust region algorithm for nonsmooth optimization", *Math. Prog.* 31 (1985) 269～285. (1985b)

[231] Y. Yuan, "An only 2-step Q-superlinear convergence example for some algorithms that use reduced Hessian approximation", *Math. Prog.* 32 (1985) 224～231. (1985c)

[232] Y. Yuan, *Some Theories and Algorithms in Nonlinear Programming* Ph.

D. Thesis, University of Cambridge, UK., 1985. (1985d)

[233] Y. Yuan, "On a subproblem of trust region algorithms for constrained optimization", *Math. Prog.* 47 (1990) 53~63. (1990a)

[234] Y. Yuan, "On self-dual update formulae in the Broyden family". Report, Computing Center, the Chinese Academy of Sciences, (1990), to appear *in Optimisation Methods and Software*. (1990b)

[235] Y. Yuan, "A dual algorithm for minimizing a quadratic function with two quadratic constraints", *J. Computational Mathematics* 9 (1991) 348 ~359.

[236] W. I. Zangwill, "Minimizing a function without calculating derivatives", *The Computer J.* 10 (1967) 293~296. (1967a)

[237] W. I. Zangwill. "Non-linear programming via penalty functions", *Management Sci.* 13 (1967) 344~358. (1967b)

[238] J. Zowe, "Nondifferentiable optimization a motivation and a short introduction in to the subgradient and the bundle concept", in: K. Schittkewski, ed., *Computational Mathematical Programming* (Springer-Verlag, Berlin, 1985) pp. 321~356.

[239] 濮定国、俞文鮆: "DFP 方法的收敛性"《曲阜师范大学学报》19 (1988), 63~69.

[240] 谢元富: "变分意义下最佳变尺度公式的研究——一个新的变尺度公式的导出",《数学学报》32 (1989), 721~736.

[241] 赵小平:《差分变尺度法的收敛性》博士论文, 1990.